

Deconvolution and Restoration of Optical Endomicroscopy Images

Ahmed Karam Eldaly, Yoann Altmann, Antonios Perperidis, Nikola Krstajić,
Tushar Choudhary, Kevin Dhaliwal, and Stephen McLaughlin

Abstract

Optical endomicroscopy (OEM) is an emerging technology platform with preclinical and clinical imaging utility. Pulmonary OEM via multicore fibres has the potential to provide *in vivo in situ* molecular signatures of disease such as infection and inflammation. However, enhancing the quality of data acquired by this technique for better visualization and subsequent analysis remains a challenging problem. Cross coupling between fiber cores is one of the main reasons of poor detection performance (i.e., inflammation, bacteria, etc.). In this work, we address the problem of deconvolution and restoration of OEM data. We propose and compare four methods, three are based on the alternating direction method of multipliers (ADMM) and one is based on Markov chain Monte Carlo (MCMC) methods. Results on both synthetic and real datasets illustrate the effectiveness of the proposed methods.

Index Terms

Optical endomicroscopy, Irregular sampling, Deconvolution, Restoration, Optimization, ADMM, Bayesian estimation, Markov Chain Monte Carlo methods.

I. INTRODUCTION

Pneumonia is a major cause of morbidity and mortality in mechanically ventilated patients in intensive care [1]. However, the accurate diagnoses and monitoring of suspected pneumonia

A. K. Eldaly, Y. Altmann, A. Perperidis and S. McLaughlin are with the Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK. (Emails: {AK577; Y.Altmann; A.Perperidis; S.Mclaughlin}@hw.ac.uk)

N. Krstajić, T.Choudhary and K. Dhaliwal are with the EPSRC IRC "Hub" in Optical Molecular Sensing & Imaging, MRC Centre for Inflammation Research, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK (Emails: {N.Krstajic; T.Choudhary; Kev.Dhaliwal}@ed.ac.uk)

Part of this work was supported by the EPSRC via grant EP/K03197X/1.

remains challenging [2]. Current methodologies consist of culturing bronchoalveolar lavage fluid (BALF) retrieved from bronchoscopy, but this often takes 48 hours to yield a result which still has low specificity and sensitivity [3]. Structural imaging with X-ray or computed tomography (CT) scans are also often non-diagnostic. An emerging methodology being developed by our group [4][5][6][7] is to deploy molecular optical imaging to delineate bacteria or inflammatory processes. Bacteria labelled with chemical SmartProbes are detectable with optical endomicroscopy (OEM) in the distal lung [5]. These SmartProbes bind to the bacteria and generate a fluorescent signal when imaged using a fluorescent OEM system. Several candidate smartprobes have been designed to target neutrophil recruitment in inflammation [6], bacterial detection [5] and fibrogenesis [7] in conjunction with OEM. OEM can be based upon widefield [4] or scanning/confocal systems. Confocal endomicroscopy provides optical sectioning allowing higher contrast imaging as tissue background fluorescence is largely suppressed [4]. Commercial confocal fiber optic imaging devices, such as the Cellvizio system developed by Mauna Kea [8][9] provide sparse and usually irregularly-spaced intensity readings of the scene, due to the irregular packing of the fiber cores. Fluorescently labelled bacteria are small relative to the size of the individual optical fibres, so individual stained bacteria are only visible by single fibres, which then tend to be brighter than their immediate neighbours [10].

One of the main challenges of OEM images is enhancing the restoration of the signals at the receiver for better image visualization and/or subsequent analysis. Fiber core cross coupling is one of the main reasons for image degradation in this type of imaging; it causes blurring in the resulting images, whose restoration is formulated as an inverse problem. We will discuss in detail cross coupling effect in Section II. In image deconvolution and restoration, a noisy observation vector \mathbf{y} (vectorized image), of vectorized image of an original intensity vector \mathbf{x} , is modelled by the following linear model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}, \quad (1)$$

where \mathbf{A} is the matrix representation of a convolution operator which can model different degradations, such as relative motion between the camera and the subject (motion blur), bad focusing (defocusing blur), fiber cores in microscopy, among others (see examples in [11]). In (1), the vector \mathbf{w} stands for additive noise modelling observation noise and model mismatch and is assumed to be a white Gaussian noise sequence. We specify the dimensions of the variables later in the text depending on the reconstruction method.

The problem of estimating \mathbf{x} from \mathbf{y} is called a linear inverse problem (LIP); for most scenarios of practical interest, this is an ill-posed LIP (IPLIP), i.e., the matrix \mathbf{A} is singular or very ill-conditioned. Consequently, this IPLIP requires additional regularization (or prior information, in Bayesian inference terms) in order to reduce uncertainties and improve estimation.

In [12]–[14], the problem of estimating \mathbf{x} , given \mathbf{y} is formulated as an unconstrained optimization problem as follows

$$\underset{\mathbf{x}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2 + \lambda \phi(\mathbf{x}) \quad (2)$$

where $\phi(\cdot)$ is the regularization function, $\|\cdot\|_2$ is the common ℓ_2 -norm, and $\lambda \in \mathbb{R}_+$ is a regularization parameter. For solving problems of the form (2), state-of-the-art algorithms potentially belonging to the iterative shrinkage/thresholding family [12]–[14] can be used. In [13], [15], the unconstrained problem in (2) is solved by an algorithm called split augmented Lagrangian shrinkage algorithm (SALSA) which is based on variable splitting [16], [17]. The idea is to transform the unconstrained problem (2) into another equivalent unconstrained one via a variable splitting trick, that can be solved by an augmented Lagrangian (AL) method [18], such as the alternating direction method of multipliers (ADMM) [19], [20].

Alternatively, many studies have considered hierarchical Bayesian models to solve the deconvolution and restoration problem [21]–[25]. These models offer a flexible and consistent methodology to deal with uncertainty in inference when limited amount of data or information is available. Moreover, other unknown parameters can be jointly estimated within the algorithm such as noise variance(s) and regularization parameters. As such, they represent an attractive way to tackle ill-posed problems. These methods rely on selecting an appropriate prior distribution for the unknown image and other unknown parameters. The full posterior can then be derived from the Bayes' rule, and then exploited by optimization or simulation-based (Markov chain Monte Carlo) methods.

The main contributions of this work are threefold:

- 1) We address the problem of double deconvolution of OEM images. To the best of our knowledge, it is the first time this problem is addressed in a statistical framework.
- 2) We develop algorithms dedicated to irregularly sampled images which do not rely on strong assumptions on the sampling patterns. The developed methods can thus be applied to a wide range of imaging systems, and multi-core fiber designs.
- 3) We develop four different methods (three optimization-based and one simulation-based)

and compare them using extensive simulations conducted using controlled and real data. The remaining sections of the paper are organized as follows. Section II formulated the problem of deconvolution and restoration of OEM data. Section III details the four proposed methods. Results of simulations conducted using synthetic and real datasets are discussed in Section IV and Section V, respectively. Conclusions and future work are finally reported in Section VI.

II. PROBLEM FORMULATION

Fig. 1 illustrates the concept of cross coupling between fiber cores. If an individual fiber core is illuminated, the neighbouring cores will be affected by a specific percentage of the incident light on the illuminated core. Experimental results in current fiber bundle (which might be different for other bundles) showed that around 60% of the light transmitted through a single core remains on that core, around 25% migrates to the immediate neighbouring cores and around 15% to the further neighbours.

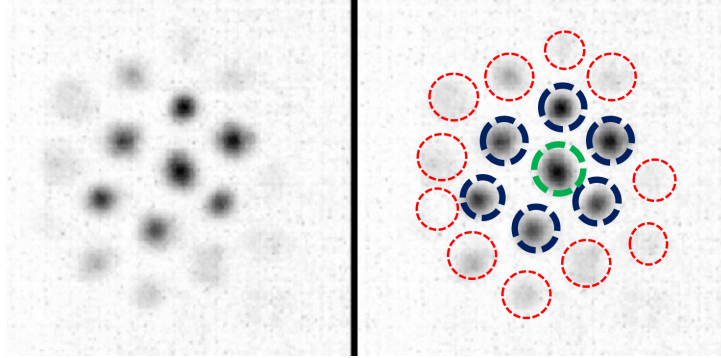


Fig. 1: Example of cross coupling between fiber cores, the green circle represents the central illuminated core and the blue and red ones represent the immediate and further neighbours respectively. For better visualization, dark (respectively bright) regions correspond to high (respectively low) intensity regions.

Fig. 2 illustrates how we construct the forward observation model to mimic the same output as our endomicroscopy imaging system. The first image of the figure represents the illumination of one fiber core. This results in cross coupling effect at the neighbouring cores (convolution with a first linear operator H_0), then the spatial blurring effect around each fiber core (convolution with a second linear operator C) and finally the fourth image of the figure shows final system output after adding white Gaussian noise.

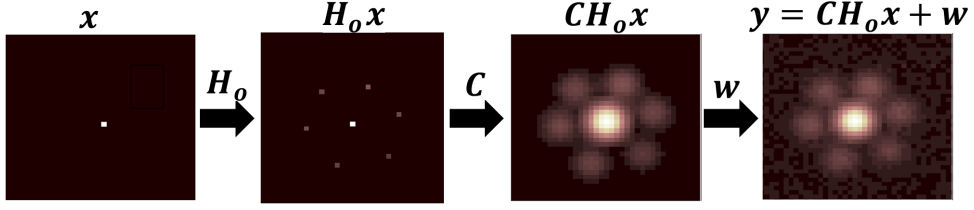


Fig. 2: Representation of the endomicroscopy system output images.

The linear model in (1) can now be written as

$$\mathbf{y} = \mathbf{C}\mathbf{H}_o\mathbf{x} + \mathbf{w} \quad (3)$$

where \mathbf{A} in (1) is replaced by $\mathbf{C}\mathbf{H}_o$ in (3), where $\mathbf{H}_o = \mathbf{D}^T\mathbf{H}\mathbf{D}$, the vector $\mathbf{y} \in \mathbb{R}^N$ is the observed vectorized data matrix (N represents the total number of pixels in the image), $\mathbf{x} \in \mathbb{R}^N$ is the image to be restored, $\mathbf{D} \in \mathbb{R}^{N_1 \times N}$ is a diagonal decimation matrix (N_1 represents number of fiber cores in the image); a value of one is assigned to the pixels representing the central fiber core locations and zero otherwise, $\mathbf{H} \in \mathbb{R}^{N_1 \times N_1}$ is the first convolution operator representing cross coupling between fiber cores, $\mathbf{D}^T \in \mathbb{R}^{N \times N_1}$ is the transposed decimation matrix, and $\mathbf{C} \in \mathbb{R}^{N \times N}$ is the second convolution operator representing spatial blurring on each core. From preliminary results, we propose to model cross-coupling by an isotropic 2D Gaussian kernel applied to the fiber intensities as follows

$$[\mathbf{H}]_{i,j} = \exp\left(\frac{-d_{i,j}^2}{2\sigma_H^2}\right) \quad (4)$$

where $d_{i,j}$ denotes the euclidean distance between the cores (or spatial locations) i and j . From (4), it can be seen that close fiber cores will be more coupled than distant ones. The value of σ_H^2 , which controls the amount of cross-coupling (the higher σ_H^2 the more coupling) and which is system dependent, is adjusted from preliminary measurements (calibration). Note that other cross-coupling models could also be considered depending on the imaging system used.

The spatial blur affecting each fiber core \mathbf{C} is also modelled by a Gaussian spatial filter, as illustrated in Fig. 3, which shows a background image of the endomicroscopy imaging system, and a zoomed-in region of this image, bright and dark areas represent fiber cores and their cladding, respectively. The intensity profile across one line in this image is a series of Gaussian kernels. The variation of their shape and width is due to the variation in core sizes.

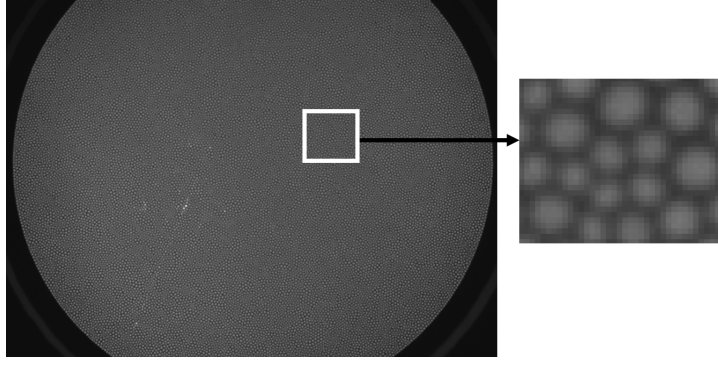


Fig. 3: A background image and a zoomed part of the image.

For simplification, a single Gaussian kernel of a specific size is used, which will simplify the model, and will reduce the computational cost as we will see in Section III-A,

The size of OEM images in our case is 1200×800 pixels and these images consist of approximately 8,000 fiber cores (or measurements). This corresponds to approximately 1.2% of the original pixels. Several approaches can be considered to process these images. Indeed, we can use as input of the deconvolution algorithm either the full observed image or (when the second convolution operator is sufficiently spatially concentrated) a single intensity per core. Considering the full image is more rigorous but potentially computationally expensive for long image sequences (videos). Similarly, one might want to reconstruct the full image directly (e.g., for visualization purposed) or only estimate the intensities at the fiber core locations. In this case the estimated full image can be recovered by interpolating the deconvolved subsampled image. Consequently, we propose four methods to solve the image restoration problem, three are optimization-based algorithms and one is a simulation-based method. The convolution operator \mathbf{A} in (1) represents the cross coupling between the fiber cores (i.e., $\mathbf{A} = \mathbf{H}_o$) if we are processing the central fiber core pixels ($\mathbf{y} \in \mathbb{R}^{N_1}$) and estimating only central fiber core intensities ($\mathbf{x} \in \mathbb{R}^{N_1}$), $\mathbf{A} = \mathbf{H}\mathbf{D}$ if we are processing central core intensities ($\mathbf{y} \in \mathbb{R}^{N_1}$) and restoring a full image ($\mathbf{x} \in \mathbb{R}^N$), and is equal to $\mathbf{A} = \mathbf{C}\mathbf{H}$ if we are processing the full image ($\mathbf{y} \in \mathbb{R}^N$) and restoring a full image ($\mathbf{x} \in \mathbb{R}^N$). Table I gives summarizes the four proposed methods in term of model and input/output parameters. Each method is named in the following way, input-output-method, where 'I' and 'C' refer to 'Image' and vector representing central fiber 'Core' intensities respectively, whereas 'o' and 's' refer to optimization and simulation based methods respectively.

	Methods			
	Ho	CIo	CCo	CCs
Model	$y = CH_{\circ}x + w$	$y = HDx + w$	$y = Hx + w$	
Input y	Full image ($y \in \mathbb{R}^N$)	A vector representing central fiber core intensities ($y \in \mathbb{R}^{N_1}$)		
Output x	Full image ($x \in \mathbb{R}^N$)		A vector representing central fiber core intensities ($x \in \mathbb{R}^{N_1}$)	

TABLE I: A summary for the four proposed methods

III. ESTIMATION STRATEGIES

A. ADMM-Based Deconvolution via Two Deconvolution Operators (Ho)

Similar to (1), we formulate the problem of deconvolving and restoring an image x from a noisy image y as follows

$$y = CD^T HDx + w \quad (5)$$

as in that case $A = CD^T HD$. Estimating x in (5) under positivity constraints (as we expect the recovered intensity image to be positive) can be formulated as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \frac{1}{2} \|CD^T HDx - y\|_2^2 + \lambda \phi(Mx), \\ \text{subject to} \quad & x \succeq 0 \end{aligned} \quad (6)$$

$\phi(Mx)$ is a regularization whose influence is controlled by λ . In this method, different regularization functions are investigated, including total variation (TV) [26], and Laplacian [27] regularizations. The Laplacian filter provides the best results in terms of root mean square error (RMSE) and visual quality of the reconstructed image. We observed that the Laplacian regularization, which promotes smoothness by minimizing the second derivatives of the image, generally yields an algorithm which converge faster than when using the TV regularization. The unconstrained version of (6) can be written as

$$\underset{x}{\text{minimize}} \quad \frac{1}{2} \|CD^T HDx - y\|_2^2 + \lambda \phi(Mx) + i_{\mathbb{R}_+}(x) \quad (7)$$

where $i_{\mathbb{R}_+}(x) = \sum_{i=1}^N i_{R_+}(x_i)$ is the indicator function which is equal to zero if x_i is positive and $+\infty$ otherwise. Given the objective function in (7), we can obtain the following

(constrained) equivalent formulation

$$\begin{aligned}
& \underset{\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7}{\text{minimize}} && \frac{1}{2} \|\mathbf{v}_1 - \mathbf{y}\|_2^2 + \lambda \phi(\mathbf{v}_6) + i_{\mathbb{R}^+}(\mathbf{v}_7) \\
& \text{subject to} && \mathbf{v}_1 = C\mathbf{v}_2, \mathbf{v}_2 = D^T \mathbf{v}_3, \mathbf{v}_3 = H\mathbf{v}_4 \\
& && \mathbf{v}_4 = D\mathbf{v}_5, \mathbf{v}_5 = \mathbf{u}, \mathbf{v}_6 = M\mathbf{u}, \mathbf{v}_7 = \mathbf{u}
\end{aligned} \tag{8}$$

The optimization problem in (8) can be written in a compact form as follows

$$\underset{\mathbf{u}, \mathbf{V}}{\text{minimize}} \quad g(\mathbf{V}), \quad \text{subject to} \quad \mathbf{G}\mathbf{u} + \mathbf{F}\mathbf{V} = \mathbf{0} \tag{9}$$

where

$$\begin{aligned}
\mathbf{V} &= (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7) \\
g(\mathbf{V}) &= \frac{1}{2} \|\mathbf{v}_1 - \mathbf{y}\|_2^2 + \lambda \phi(\mathbf{v}_6) + i_{\mathbb{R}^+}(\mathbf{v}_7) \\
\mathbf{G} &= \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \\ \mathbf{M} \\ \mathbf{I} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} -\mathbf{I} & \mathbf{C} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & D^T & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & H & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & D & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}
\end{aligned}$$

The ADMM algorithm for solving (9) is shown in Algorithm 1 where: $\mathcal{L}(\mathbf{u}, \mathbf{V}, \mathbf{D}_L) = g(\mathbf{u}, \mathbf{V}) + \frac{\mu}{2} \|\mathbf{G}\mathbf{u} + \mathbf{F}\mathbf{V} - \mathbf{D}_L\|_2^2$ is the augmented Lagrangian associated with (9), $\mu > 0$ is a positive parameter which is updated to keep the primal and dual residual norms within a factor of 10 of one another [28], and $\frac{\mathbf{D}_L}{\mu}$ denotes the Lagrange multipliers associated with the constraint $\mathbf{G}\mathbf{u} + \mathbf{F}\mathbf{V} = \mathbf{0}$. During each step of the iterative algorithm, \mathcal{L} is optimized with respect to (w.r.t.) \mathbf{u} (step 3) and \mathbf{V} (step 4) and then the Lagrange multipliers are updated (step 5).

The matrix \mathbf{G} is full column rank, and the function $g(\cdot)$ introduced in (9) is closed, proper, and convex. Then, *Theorem 1* of [20] ensures that, for any $\mu > 0$, if (9) has a solution, say \mathbf{u}^* , then the sequence $\{\mathbf{u}^{(k)}\}$, converges to \mathbf{u}^* . If (9) does not have a solution, then at least one of the sequences $\{\mathbf{u}^{(k)}\}$ or $\{\mathbf{D}_L^{(k)}\}$ diverges. The stopping criterion we use is $\|\mathbf{G}\mathbf{u}^{(k)} + \mathbf{F}\mathbf{V}^{(k)}\|_F \leq \epsilon$, where $\epsilon = \sqrt{N} \times 10^{-5}$ [29].

The algorithmic details for solving (9) are provided in the Appendix.

Algorithm 1 Ilo Algorithm

- 1: set $k = 0$, choose $\mu > 0, \mathbf{u}^{(0)}, \mathbf{V}^{(0)}, \mathbf{D}_L^{(0)}$
 - 2: **repeat**
 - 3: $\mathbf{u}^{(k+1)} \leftarrow \arg \min_{\mathbf{u}} \mathcal{L}(\mathbf{u}, \mathbf{V}^{(k)}, \mathbf{D}_L^{(k)})$
 - 4: $\mathbf{V}^{(k+1)} \leftarrow \arg \min_{\mathbf{V}} \mathcal{L}(\mathbf{u}^{(k+1)}, \mathbf{V}, \mathbf{D}_L^{(k)})$
 - 5: $\mathbf{D}_L^{(k+1)} \leftarrow \mathbf{D}_L^{(k)} - (\mathbf{G}\mathbf{u}^{(k+1)} + \mathbf{F}\mathbf{V}^{(k+1)})$
 - 6: **Update** μ [28].
 - 7: **until** some stopping criterion is satisfied.
-

B. ADMM-Based Deconvolution via One Deconvolution Operator (CIo)

Although the model used in the Ilo method is more rigorous, it is potentially computationally expensive for long image sequences since it considers full images as input and the output (see Appendix for details). In this method, the second convolution operator \mathbf{C} is omitted, we can then summarize the information of each core by a single intensity value. Three ways can be used, either integrating over the intensities of each core, or over a small radius, which results in a reduction in signal to noise ratio (SNR) and/or errors due to the overlapping between the cores, or using central pixel of each fiber core. The later way provides stable representation of each core's intensity since it is less affected by noise and not affected by cores overlap. Following this strategy, the model now becomes

$$\mathbf{y} = \mathbf{H}\mathbf{D}\mathbf{x} + \mathbf{w} \quad (10)$$

where $\mathbf{y} \in \mathbb{R}^{N_1}$, which means that it is a vector representing central fiber core intensities, and $\mathbf{x} \in \mathbb{R}^N$, which means that a full image is restored.

In a similar fashion to the Ilo method, we can estimate \mathbf{x} in (10) by solving a convex optimization problem using an ADMM-based algorithm. Note that since we consider $\mathbf{H}\mathbf{D}$ instead of $\mathbf{C}\mathbf{D}^T\mathbf{H}\mathbf{D}$, we can use a simpler splitting scheme and thus solve a minimization problem w.r.t. a reduced number of unknown (and in practice with a less computationally demanding algorithm). For this method, referred to as CIo, we used the same Laplacian-based regularization as in Ilo.

C. ADMM-Based Deconvolution via One Deconvolution Operator (CCo)

The Ilo, and the CIo methods restore a full image which is important for visualization, but is computationally expensive. Moreover, post processing like bacteria detection for instance

can only be carried out in core's central pixels, which means that the computation cost can be reduced by only having central core intensities, which are much less than total number of pixels in the image, at both input and output of an algorithm. Hence, the convergence can generally be improved since less number of parameters are required to be estimated. Furthermore, a full image for visualization purposes can be obtained by linear or non-linear interpolation of the deconvolved core intensities. In this method, referred to as CCo the input $\mathbf{x} \in \mathbb{R}^{N_1}$ and the output $\mathbf{y} \in \mathbb{R}^{N_1}$ are two vectors representing the central core intensities, where $N_1 \ll N$. The model now becomes

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \quad (11)$$

In a similar fashion to the IIo method, we can estimate \mathbf{x} in (11) by solving a convex optimization problem using an ADMM-based algorithm. Using a variable splitting scheme in this method results in only a single splitting step since we consider \mathbf{H} instead of $\mathbf{C}\mathbf{D}^T\mathbf{H}\mathbf{D}$ in the IIo method and $\mathbf{H}\mathbf{D}$ in the CIo method, and thus solve a minimization problem w.r.t. just two unknowns.

The regularization function $\phi(\mathbf{x})$ in this method is (up to an additive constant) the negative logarithm of the output image prior considered in the CCs method, e.g., $\phi(\mathbf{x}) = \frac{\mathbf{x}^T \mathbf{\Delta} \mathbf{x}}{2}$ where the matrix $\mathbf{\Delta}$ is used to correlate the pixels spatially, and will be specified later in the text.

D. MCMC Based Deconvolution via One Deconvolution Operator (CCs)

As mentioned in the introduction, one main disadvantage of the ADMM based methods is the difficulty in choosing the regularization parameter λ . This problem can be overcome by an MCMC method which requires less user supervision. Furthermore, the noise variance can be jointly estimated within the algorithm.

Similarly to the CCo method, the input $\mathbf{x} \in \mathbb{R}^{N_1}$ and the output $\mathbf{y} \in \mathbb{R}^{N_1}$ are two vectors representing central fiber core intensities. The noise \mathbf{w} is assumed to be additive white noise which is assumed to be an independent and identically distributed (i.i.d) zero mean Gaussian noise with variance σ^2 , denoted as $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$, where \sim means “is distributed according to” and \mathbf{I} is the identity matrix.

1) Hierarchical Bayesian Model: This section introduces a hierarchical Bayesian model proposed to estimate the unknown parameter vector \mathbf{x} . This model is based on the likelihood function of the observations and on prior distributions assigned to the unknown parameters. Likelihood: (11) shows that $\mathbf{y} | (\mathbf{x}, \sigma^2) \sim \mathcal{N}(\mathbf{H}\mathbf{x}, \sigma^2 \mathbf{I})$. Consequently, the likelihood function

of \mathbf{y} can be expressed as

$$f(\mathbf{y}|\mathbf{x}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left(-\frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2}{2\sigma^2}\right) \quad (12)$$

Prior for the underlying intensity field \mathbf{x} : For many applications, the intensity values of the scene to be recovered are likely to be spatially correlated. An interesting way to take possibly correlated intensities is to consider Markov random fields (MRF) to build a prior for \mathbf{x} . MRFs assume that the distribution of a given intensity x_n conditionally to the other intensity values of the image equals the distribution of this parameter conditionally to its spatial neighbours, i.e., $f(x_n|\mathbf{x}_{\setminus x_n}) = f(x_n|\mathbf{x}_{\mathcal{V}_n})$, where \mathcal{V}_n is the index set of the neighbours of x_n , $\mathbf{x}_{\setminus x_n}$ denotes the vector \mathbf{x} whose element x_n has been removed, and $\mathbf{x}_{\mathcal{V}_n}$ is the subset of \mathbf{x} composed of the elements whose indexes belong to \mathcal{V}_n . In this work, we specify $f(x_n|\mathbf{x}_{\mathcal{V}_n})$ as

$$f(x_n|\mathbf{x}_{\mathcal{V}_n}, \gamma^2) \propto \exp\left(-\frac{1}{\gamma^2} \sum_{n' \in \mathcal{V}_n} \frac{(x_n - x_{n'})^2}{d_{n,n'}}\right) 1_{\mathbb{R}^+}(\mathbf{x}) \quad (13)$$

where $d_{n,n'}$ denotes the distance between the spatial locations n, n' and γ^2 controls the global correlation between intensities, and $1_{\mathbb{R}^+}(\mathbf{x})$ is the indicator function defined on the positive set of \mathbf{x} . Equation (13) promotes smooth intensity variations between neighbours while ensuring that the prior dependence between neighbours decrease as $d_{n,n'}$ increases. In this work $d_{n,n'}$ is the standard euclidean distance, but other distances could be used (e.g., distances on manifolds). It can be shown that the resulting joint prior $f(\mathbf{x}|\gamma^2)$ can be expressed as

$$f(\mathbf{x}|\gamma^2) \propto (\gamma^2)^{-d/2} \exp\left(-\frac{\mathbf{x}^T \Delta \mathbf{x}}{2\gamma^2}\right) 1_{\mathbb{R}^+}(\mathbf{x}) \quad (14)$$

where

$$[\Delta]_{i,j} = \begin{cases} \sum_{j \in \mathcal{V}_i} 1/d_{i,j} & \text{if } i = j \\ 0 & \text{if } j \notin \mathcal{V}_i \\ -1/d_{i,j} & \text{else} \end{cases} \quad (15)$$

and $d = \text{rank}(\Delta)$.

Prior for the noise variance σ^2 : A conjugate inverse-Gamma \mathcal{IG} prior is assigned to the noise variance σ^2

$$f(\sigma^2|\alpha, \beta) \sim \mathcal{IG}(\alpha, \beta) \quad (16)$$

where $\alpha = 10$ is fixed, while the hyperparameter β is jointly estimated within the algorithm.

Prior for the hyperparameter β : The hyperparameter associated with the parameter prior defined above is assigned to a Gamma distribution:

$$\beta \sim \mathcal{G}(\alpha_0, \beta_0) \quad (17)$$

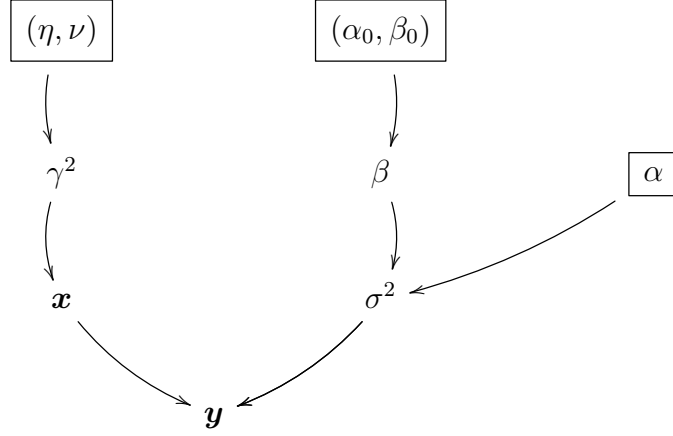


Fig. 4: Graphical model for the proposed hierarchical Bayesian model (fixed quantities appear in boxes).

where α_0 and β_0 are fixed and user-defined parameters which might depend on the dynamics of the image to be recovered. In this work, we fixed $(\alpha_0, \beta_0) = (10, 0.1)$.

Prior for the hyperparameter γ^2 : To reflect the lack of prior knowledge about the regularization parameter γ^2 in (14), the following weakly informative inverse-Gamma prior is assigned to it.

$$\gamma^2 \sim \mathcal{IG}(\eta, \nu) \quad (18)$$

where (η, ν) are fixed to $(\eta, \nu) = (10^{-3}, 10^{-3})$. The next section derives the joint posterior distribution of the unknown parameters associated with the proposed Bayesian model and studies an MCMC methods to sample from this posterior. These samples can then be used to approximate appropriate Bayesian estimators.

Joint posterior distribution: Assuming the parameters \mathbf{x} and σ^2 are a priori independent, the joint posterior of the parameter vector $\boldsymbol{\theta} = \{\mathbf{x}, \sigma^2\}$ and hyperparameters $\boldsymbol{\phi} = \{\beta, \gamma^2\}$ can be expressed as

$$f(\boldsymbol{\theta}, \boldsymbol{\phi} | \mathbf{y}) \propto f(\mathbf{y} | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \boldsymbol{\phi}) f(\boldsymbol{\phi}) \quad (19)$$

where

$$\begin{aligned} f(\boldsymbol{\theta} | \boldsymbol{\phi}) &= f(\mathbf{x} | \gamma^2) f(\sigma^2 | \beta) \\ f(\boldsymbol{\phi}) &= f(\gamma^2) f(\beta) \end{aligned} \quad (20)$$

The directed acyclic graph (DAG) summarizing the structure of proposed Bayesian model is depicted in Fig. 4.

2) *Bayesian Inference*: To overcome the challenging derivation of Bayesian estimators associated with $f(\boldsymbol{\theta}, \boldsymbol{\phi}|\mathbf{y})$, we propose to use an efficient MCMC method to generate samples asymptotically distributed according to the posterior presented in (19). More precisely, we consider a Gibbs sampler described in the next part of this section. The principle of the Gibbs sampler is to sample according to the conditional distributions of the posterior of interest [[30], Chap. 10]. In this work, we propose to sample sequentially the elements of $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ using moves that are detailed below.

Sampling the intensity field \mathbf{x} : From (19), since the prior (14) is conjugate to the Gaussian distribution, the full conditional distribution of \mathbf{x} is given by

$$f(\mathbf{x}|\mathbf{y}, \sigma^2) \sim \mathcal{N}_{\mathbb{R}^+}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (21)$$

where

$$\begin{aligned} \boldsymbol{\mu} &= (\sigma^{-2} \mathbf{y}^T \mathbf{H} \boldsymbol{\Sigma})^T \\ \boldsymbol{\Sigma} &= (\sigma^{-2} \mathbf{H}^T \mathbf{H} + \gamma^{-2} \boldsymbol{\Delta})^{-1} \end{aligned} \quad (22)$$

Sampling from (21) can be achieved efficiently by using the Hamiltonian method proposed in [31].

Sampling the noise variance σ^2 : By cancelling out the terms that don't depend on σ^2 from the posterior in (19), its conditional distribution can be written as

$$f(\sigma^2|\mathbf{y}, \mathbf{x}) \sim \mathcal{IG} \left(\alpha + \frac{N}{2}, \beta + \frac{\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2}{2} \right) \quad (23)$$

Sampling the hyperparameter β : It can be easily shown that β can be sampled from the following Gamma distribution

$$f(\beta|\sigma^2) \sim \mathcal{G} \left(\alpha + \alpha_0, \frac{\sigma^2 \beta_0}{\sigma^2 + \beta_0} \right) \quad (24)$$

Sampling the hyperparameter γ^2 : In a similar fashion to the noise variance, γ^2 can be sampled from the following inverse-Gamma distribution

$$f(\gamma^2|\mathbf{x}) \sim \mathcal{IG} \left(\eta + \frac{d}{2}, \nu + \frac{\mathbf{x}^T \boldsymbol{\Delta} \mathbf{x}}{2} \right) \quad (25)$$

The algorithm for generating samples asymptotically distributed according to the posterior using Gibbs sampler is shown in Algorithm 2.

The posterior mean or minimum mean square error (MMSE) estimator of \mathbf{x} can be approximated by

$$\hat{\mathbf{x}} = \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{t=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{x}^{(t)} \quad (26)$$

Algorithm 2 Gibbs Sampling Algorithm for CCs

- 1: **Fixed input parameters:** Number of burn-in iterations N_{bi} , total number of iterations N_{MC}
 - 2: **Initializations** ($k = 0$)
 - Set $\mathbf{x}^{(0)}, \sigma^{2(0)}, \beta^{(0)}, \gamma^{2(0)}$
 - 3: **Repeat** ($1 \leq k \leq N_{\text{MC}}$)
 - Sample $\mathbf{x}^{(k)}$ from (21) - Truncated MVG
 - Sample $\sigma^{2(k)}$ from (23) - \mathcal{IG}
 - Sample $\beta^{(k)}$ from (24) - \mathcal{G}
 - Sample $\gamma^{2(k)}$ from (25) - \mathcal{IG}
 - 4: **Set** $k = k + 1$.
-

where the samples from the first N_{bi} iterations (corresponding to the transient regime or burn-in period, which is determined visually from preliminary runs) of the sampler are discarded.

IV. SIMULATIONS USING SYNTHETIC DATA

The performance of the proposed methods is investigated by reconstructing a standard test image. A subsampled version of this image is obtained by considering the sampling pattern of an actual endomicroscopy system, as illustrated in Fig. 5. This figure provides an example of a homogeneous region imaged through a PROTEUS fiber bundle. Such image is used for calibration and to identify the number and positions of the fiber cores.

Fig. 6 shows the original Lena image and an example of system output after applying the model in Eq. (3). This image is created by number of samples corresponding to 1.4% of the original Lena image, in which the samples correspond to central core pixels of an actual endomicroscopy system, as illustrated in Fig. 5(b). The variance of the Gaussian kernel representing the spatial blur σ_C^2 is fixed to $\sigma_C^2 = 2.2$ in all the simulations.

We have applied the proposed methods to reconstruct the full image (as in Fig. 6(a)) from the convolved and noisy measurements (either the full image as in Fig. 6(b) or from the measured intensities at the core locations). For the methods estimating only the core intensities, a linear interpolation is then applied to recover the full images.

The performance discriminator adopted in this work to measure the quality of the deconvolved fiber cores is the root mean square error (RMSE), which is computed using intensities

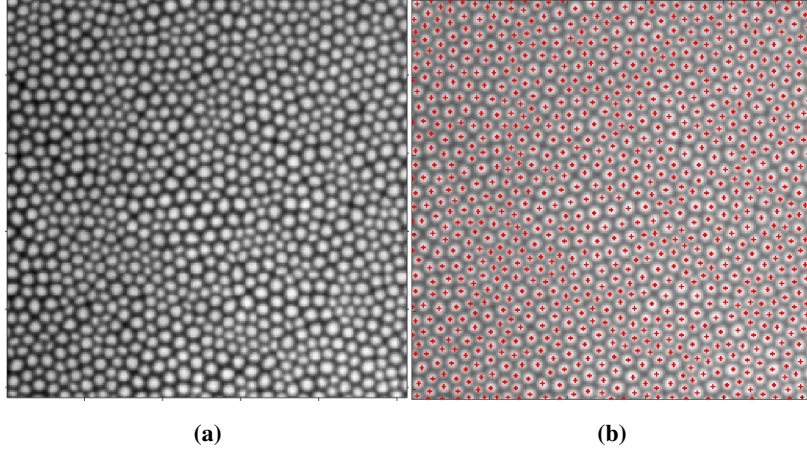


Fig. 5: (a) Example of 256×256 pixels image of the endomicroscopy system (b) Image with detected fiber core centres superimposed (red crosses).

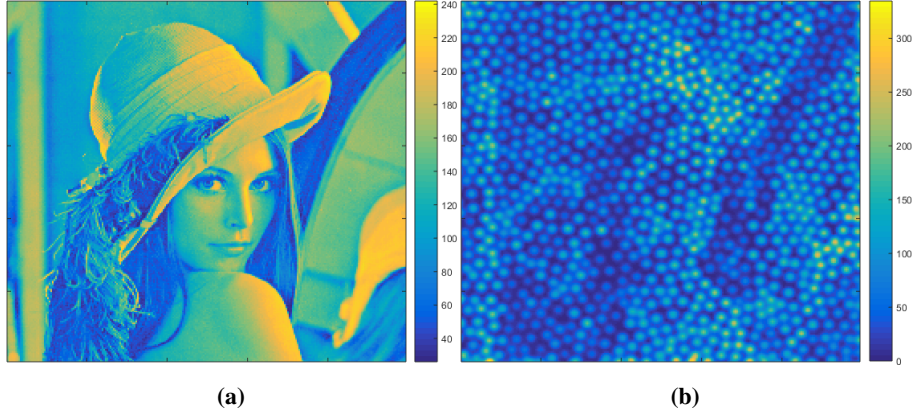


Fig. 6: Creation of the synthetic data: (a) Original image (b) example of final system output with $\sigma_H^2 = 10$ and $\sigma_N^2 = 2$.

at the core locations using

$$\text{RMSE}(\mathbf{x}, \mathbf{x}') = \sqrt{\frac{\sum_{n=1}^{N_1} [\mathbf{x}(n) - \mathbf{x}'(n)]^2}{N_1}} \quad (27)$$

where \mathbf{x} and \mathbf{x}' are the subsampled reference Lena image and its deconvolved version respectively.

In the following subsections, we present the results of the IIo, and CIo methods. The results of the CCo, and CCs methods are fairly similar to those of the CIo method.

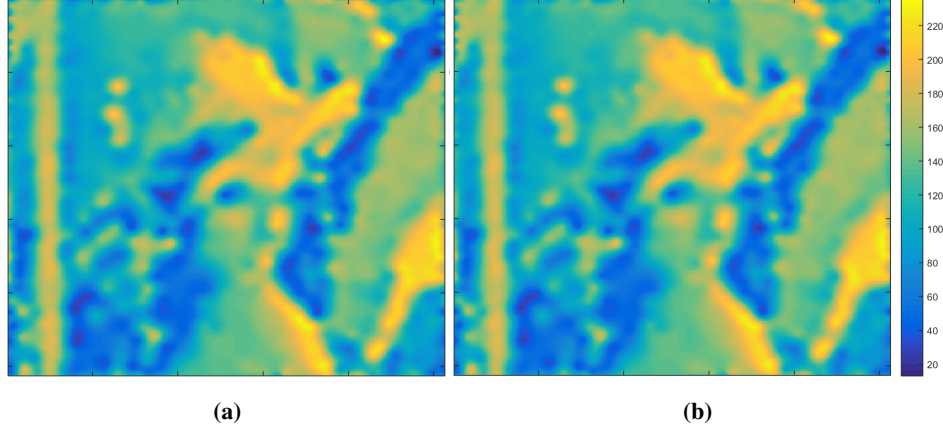


Fig. 7: Examples of output images of the Ilo algorithm (a) $\sigma_N^2 = 0$ and $\sigma_H^2 = 1$, and (b) $\sigma_N^2 = 10$ and $\sigma_H^2 = 20$.

A. ADMM Based Deconvolution via Two Deconvolution Operators (Ilo)

In this method, Different values for the regularization parameter λ are tested, we show the results corresponding to λ associated with the minimum RMSE. The input and the output of the algorithm are full images. Fig. 7 shows examples of deconvolution results in the noise-free case ($\sigma_N^2 = 0$) and noisy case ($\sigma_N^2 = 10$) and different values of σ_H^2 , we can observe that the structure of the Lena image can be recovered. Furthermore, we found that the RMSE of the algorithm output images is fairly constant as σ_H^2 increases while keeping σ_N^2 constant or vice versa.

In Fig. 8, we show a plot of RMSE (in log-scale) before and after deconvolution between the deconvolved central fiber core intensities and the original central fiber core intensities from the Lena image, versus σ_H^2 at $\sigma_N^2 = 10$. We can observe that the method is very effective since the RMSE after deconvolution is always much lower than that before deconvolution. Moreover, the gain increases with cross coupling.

Fig. 9 shows plots of RMSEs between the deconvolved subsampled image and the original subsampled Lena image for different σ_N^2 at fixed σ_H^2 in Fig. 9(a) and vice versa in Fig. 9(b). In Fig. 9(a), we can observe that the behaviour at $\sigma_H^2 = 1, 5, 10$ and 15 is almost the same. In Fig. 9(b), we can observe that RMSE is fairly constant as σ_H^2 increases at constant σ_N^2 . Furthermore, it starts to increase as σ_N^2 increases but still remains constant when changing σ_H^2 .

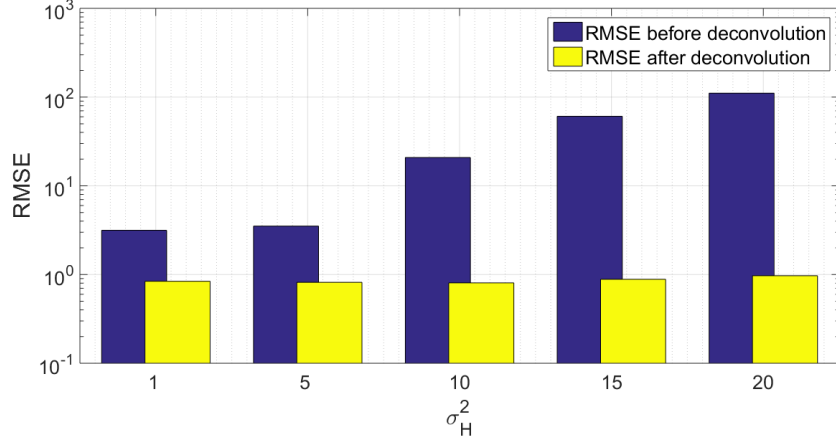


Fig. 8: Plot of RMSEs before and after deconvolution (in-log scale) versus σ_H^2 at $\sigma_N^2 = 10$ for the Ilo method.

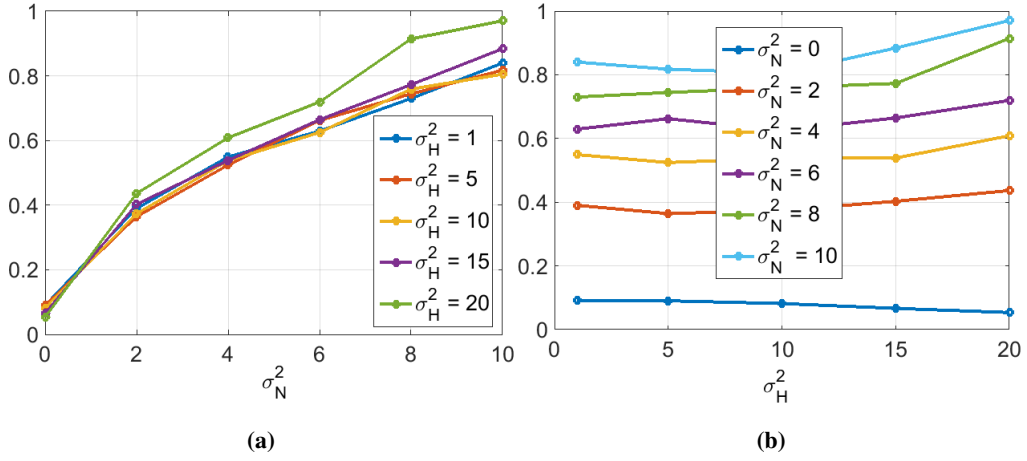


Fig. 9: Plot of RMSEs between the deconvolved subsampled image and the original subsampled Lena image versus (a) σ_N^2 at fixed σ_H^2 , and (b) σ_H^2 at fixed σ_N^2 for the Ilo method

B. ADMM Based Deconvolution via One Deconvolution Operator (Cio)

In this method, the Laplacian filter also provides the best results in terms of the quality of the reconstructed image and the RMSE in comparison with TV and DCT. The input to the algorithm is a vector representing the central fiber core intensities and the algorithm restores a full image. The deconvolved images via this method are similar to those obtained by the Ilo method shown in Fig. 7.

The RMSE of the algorithm output images are similar to those obtained by the Ilo method at which it is roughly constant as σ_N^2 increases at constant σ_H^2 or vice versa.

In Fig. 10, we show a plot of RMSE (in log-scale) before and after deconvolution between the deconvolved central fiber core and the original central fiber core intensities from the

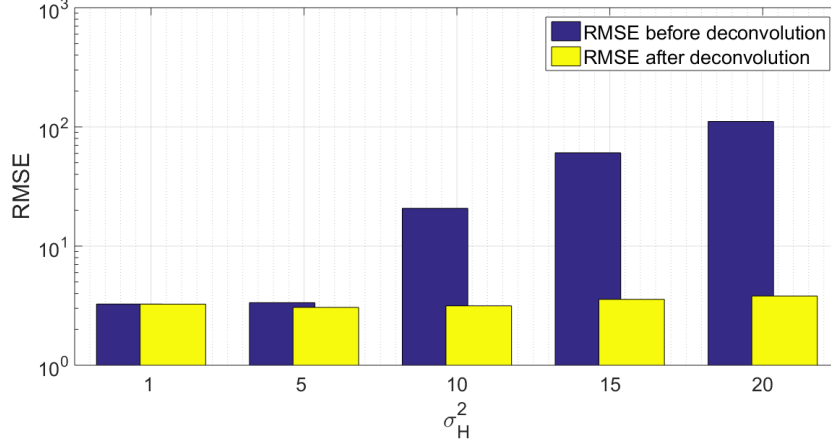


Fig. 10: Plot of RMSEs before and after deconvolution (in log-scale) versus σ_H^2 at $\sigma_N^2 = 10$ for the CIo method.

Lena image, versus σ_H^2 at $\sigma_N^2 = 10$. Still the method is effective in the deconvolution as RMSE after deconvolution is always lower than that before deconvolution, except if there is no cross coupling between the cores ($\sigma_H^2 = 1$), RMSE before and after deconvolution is equal. Moreover, the gain increases when cross coupling increases as in the IIo method.

Fig. 11 shows plots of RMSE between the deconvolved subsampled image and the original subsampled Lena image versus σ_N^2 at fixed σ_H^2 and vice versa. In Fig. 11(a), we can observe that as σ_N^2 increases, RMSE increases. Moreover, the behaviour at $\sigma_H^2 = 1, 5$ and 10 is roughly similar. In Fig. 11(b), we can observe that the RMSE is roughly constant as σ_H^2 increases at constant σ_N^2 , and starts to increase as σ_N^2 increases but still remains constant as σ_H^2 increases.

The input and the output of the CCo, and CCs methods are vectors representing the central core intensities. The full image can then be recovered by linear or non-linear interpolation of the deconvolved core intensities. For the CCs method, we used $N_{MC} = 1500$ iterations (including $N_{bi} = 500$) is used. The results obtained with these two methods are fairly similar to those obtained by CIo. Thus the results are not included here for brevity issues.

C. Comparison

In this section, we compare the four proposed methods for deconvolution and restoration of OEM images. The comparison is conducted in terms of RMSE between the central fiber cores before and after deconvolution, as well as in terms of computation time.

Fig. 12 compares RMSEs between the subsampled images before and after deconvolution

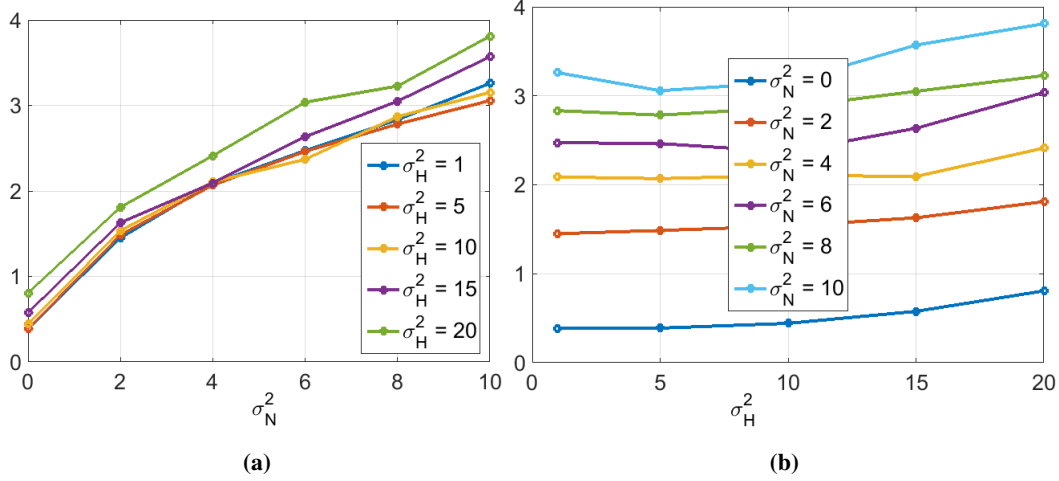


Fig. 11: Plot of RMSEs between the central fiber core intensities before and after deconvolution (a) versus σ_N^2 at fixed σ_H^2 , and (b) versus σ_H^2 at fixed σ_N^2 for the ClO method.

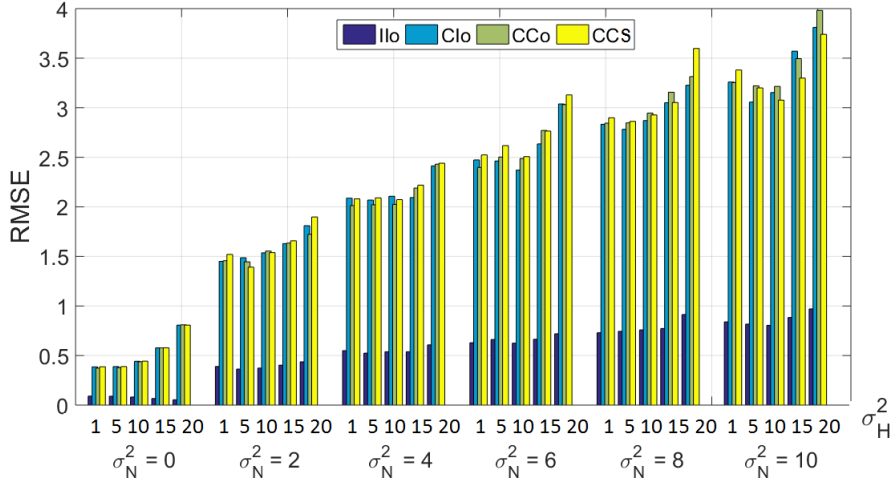


Fig. 12: Plot of RMSEs between the subsampled original Lena image and the deconvolved fiber cores by the four methods versus σ_N^2 as well as σ_H^2 .

versus different σ_N^2 as well as different σ_H^2 . We can observe that for all of the methods, as σ_N^2 increases at constant σ_H^2 , RMSE increases. On the other hand, at fixed σ_N^2 , RMSE seems to be roughly constant for $\sigma_H^2 = 1, 5$, and 10, then, it starts to increase as σ_H^2 increases. It is clear that the best method in terms of RMSE is the Ilo. In this method, RMSE is always lower than one for all of the tested σ_N^2 as well as all of the tested σ_H^2 . The behaviours of the rest of the methods remains similar.

Table II shows the average computation time (in seconds) of the four proposed methods. The experiments were conducted on ACER core-i3-2.0 GHz processor laptop with 8 GB

RAMS. Although the IIo method is the best in terms of RMSE, it is the most computationally expensive method. This method involves the successive optimization of eight objective functions with respect to high dimensional variables at each iteration as shown in the appendix. The CIo method comes on the second, then the CCs, and lastly, the CCo, at which the computation time is fractions of a second. So, there is a trade off between the deconvolution and restoration quality and the computation time.

There are two main reasons for the slow convergence of the ADMM-based algorithm (IIo and CIo), the data size either at the input or the output of the algorithms, where many pixels need to be estimated and the nature of the problem, since it is still an ill-posed problem.

Despite the relatively high computation time of the CCs method, it is a parameter free method compared to the ADMM based methods for which the regularization parameter λ should be chosen carefully.

Since the model $\mathbf{y} = \mathbf{CH}_o\mathbf{x} + \mathbf{w}$ is considered as the best in terms of RMSE, but the worst in term of computation time, other methods can be considered, instead of the ADMM method, to reduce the computation time (i.e., gradient descent, primal dual [32], etc.).

Method	IIo	CIo	CCo	CCs
Computation time	707.14	423.83	0.53	256.39

TABLE II: The average computation time (in seconds) of the four proposed methods.

D. Robustness

To test the robustness of the proposed methods, we create the data using a specific σ_H^2 and we deconvolve using different values. Following this strategy, we create the data using $\sigma_H^2 = 10$ and we deconvolve using $\sigma_H^2 = 6, 8, 12$, and 14 including $\sigma_H^2 = 10$. Due to the similarity of the results of all the proposed methods, we present the results of the IIo method and a comparison between the four methods in terms of RMSE.

Fig. 13 shows plots of RMSE between the central fiber cores in the original Lena image and the deconvolved central fiber cores versus σ_N^2 at fixed σ_H^2 and vice versa. In Fig. 13(a), we can observe that the noise variance has no effect on the deconvolution as RMSE is constant at fixed σ_H^2 . In Fig. 13(b), there is a linear relationship between RMSE and σ_H^2 at constant σ_N^2 . Furthermore, lower values of σ_H^2 than the one we created the data with (i.e., $\sigma_H^2 = 6$ and 8) yield lower RMSE than higher ones (i.e., $\sigma_H^2 = 12$ and 14). In other words, it is slightly better to underestimate σ_H^2 than to overestimate it.

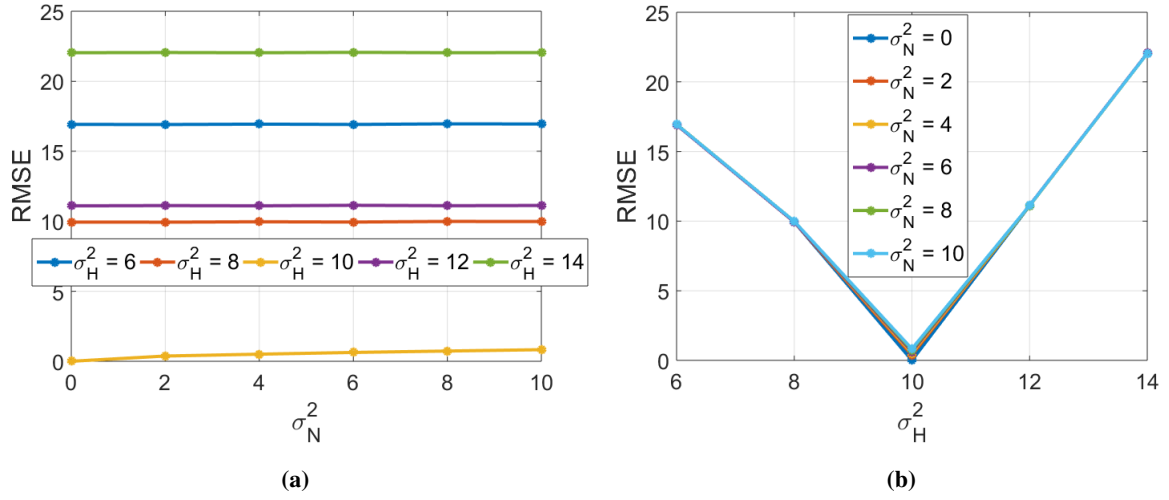


Fig. 13: Plots of RMSEs between the central fiber cores in the original Lena image and the deconvolved central fiber cores versus (a) σ_N^2 at fixed σ_H^2 , and (b) σ_H^2 at fixed σ_N^2 of the Ilo method.

We observe that deconvolution using the value we created the data with ($\sigma_H^2 = 10$) yields the minimum RMSE. Moreover, RMSE after deconvolution is always lower than that before deconvolution except for $\sigma_H^2 = 14$ at which it is higher.

Fig. 14 shows a comparison between the four proposed methods in terms of RMSE. Excluding $\sigma_H^2 = 10$; the one we created the data with; all the methods are showing similar results.

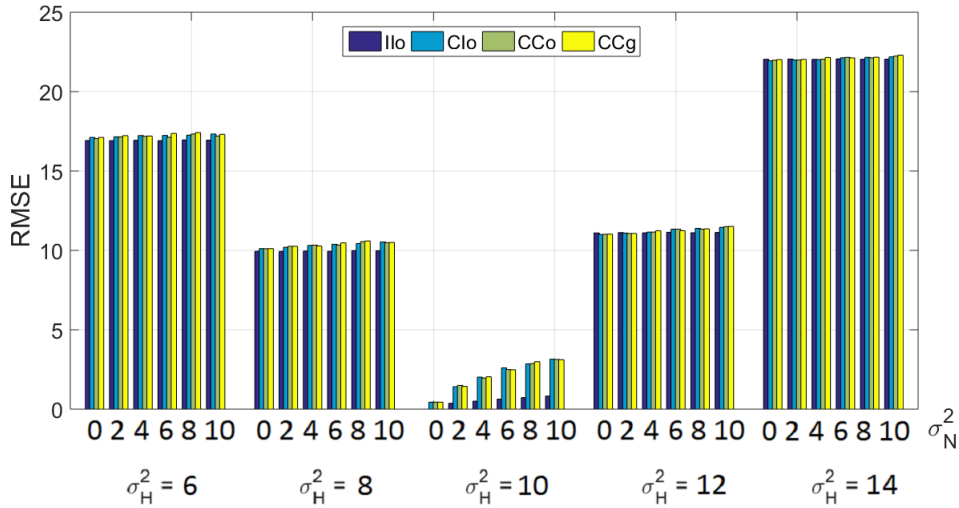


Fig. 14: Plot of RMSEs between the subsampled original Lena image and the deconvolved fiber cores by the four methods versus different σ_N^2 as well as different σ_H^2 .

V. SIMULATIONS USING REAL DATA

The performance of the proposed methods has been evaluated on two real datasets; the 1951 USAF resolution test chart and Ex-vivo human lung tissues. Plots of reconstruction error $RE = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ versus regularization parameter λ are considered for the Ilo, CIo, and CCo methods, since ground truth is not available for the data. A 2D generalized Gaussian distribution with zero mean is considered for the cross coupling convolution operator \mathbf{H} , e.g.,

$$[\mathbf{H}]_{i,j} = \exp\left(-\left(\frac{|d_{i,j}|}{\alpha}\right)^\beta\right) \quad (28)$$

where, from preliminary analysis, α and β are fixed to 4 and 0.8 respectively. Linear interpolation of deconvolved cores are considered for the CCo and CCs methods, which is compared with non-linear interpolation (based on Gaussian process regression 'GPR') [33], since it can provide confidence intervals for each interpolated pixel.

For the CCo and CCs methods, more general conjugate priors, such Gaussian Process (GP) based priors, could be used. A classical choice consists of considering a zero-mean GP with covariance matrix Δ defined as

$$\Delta(z_i, z_j) = \sigma_f^2 \exp\left(-\frac{d^2(z_i, z_j)}{2\ell^2}\right) \quad (29)$$

where $d(z_i, z_j)$ denotes the euclidean distance between the cores (or spatial locations) z_i and z_j , and σ_f^2 is the maximum allowable covariance. Note that z_i here is a 2-dimensional vector representing the i th fibre core location in the image. If $d(z_i, z_j)$ is very small, then $\Delta(z_i, z_j)$ approaches its maximum σ_f^2 . Now if z_i is distant from z_j , we have instead $\Delta(z_i, z_j) \approx 0$, i.e. the two points are considered to be a priori independent. So, for example, during interpolation at new z_* location, distant cores will have negligible effect. The amount of spatial correlation depends on the length parameter, ℓ . The values of σ_f^2 and ℓ^2 are optimized by maximum likelihood estimation of the vector to be interpolated [33]. Precisely, once the core intensities have been estimated using CCo or CCs, σ_f^2 and ℓ^2 are obtained by maximizing

$$-\frac{1}{2} \log(|\Delta(\mathbf{z}, \mathbf{z})|) - \frac{1}{2} \mathbf{x}^T (\Delta(\mathbf{z}, \mathbf{z}))^{-1} \mathbf{x}, \quad (30)$$

where $\Delta(\mathbf{z}, \mathbf{z}) \in \mathbb{R}^{N_1 \times N_1}$, $\mathbf{z} = [z_1, \dots, z_{N_1}]^T$ contains all the positions of all the observed cores (whose estimated intensities are gathered in \mathbf{x}). If we now consider a new spatial location z_* for which we want to predict the intensity x_* , the GP can be extended as follows

$$\begin{bmatrix} \mathbf{x} \\ x_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \Delta(\mathbf{z}, \mathbf{z}) & \Delta(\mathbf{z}, z_*) \\ \Delta(z_*, \mathbf{z}) & \sigma_f^2 \end{bmatrix}\right), \quad (31)$$

where $\Delta(\mathbf{z}, \mathbf{z}_*) = \Delta(\mathbf{z}_*, \mathbf{z})^T \in \mathbb{R}^{N_1}$. Eq. (31) shows that the conditional distribution of each predicted intensity given the previously estimated intensities (output of CCo or CCs), follows a Gaussian distribution $x_* | \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ whose mean and variance are given by

$$\begin{aligned} \boldsymbol{\mu} &= \Delta(\mathbf{z}_*, \mathbf{z}) \Delta(\mathbf{z}, \mathbf{z})^{-1} \mathbf{x}, \\ \boldsymbol{\Sigma} &= \sigma_f^2 - \Delta(\mathbf{z}_*, \mathbf{z}) \Delta(\mathbf{z}, \mathbf{z})^{-1} \Delta(\mathbf{z}, \mathbf{z}_*). \end{aligned} \quad (32)$$

The mean in (32) is finally used to estimate each interpolated intensity, while the variance is used to provide additional information (uncertainty) about the interpolated intensity values.

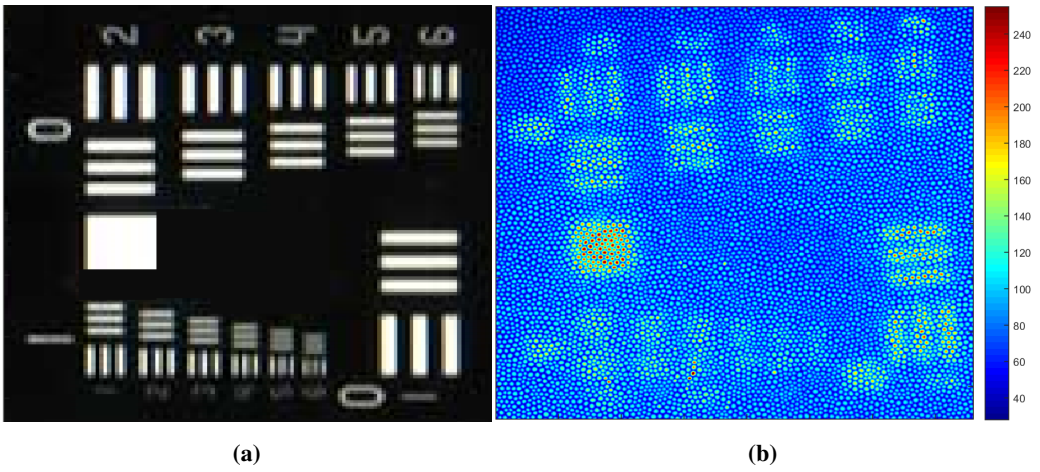


Fig. 15: (a) Scanned image of an USAF 1951 Resolution test chart. (b) The 1951 USAF resolution test chart imaged by our OEM system.

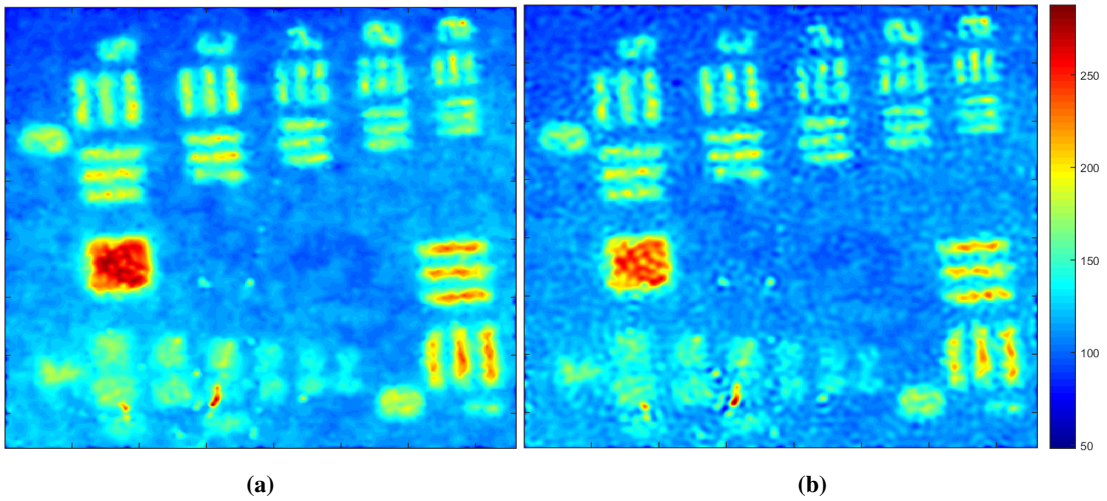


Fig. 16: (a) Linear, and (b) non-linear interpolation of central cores before deconvolution.

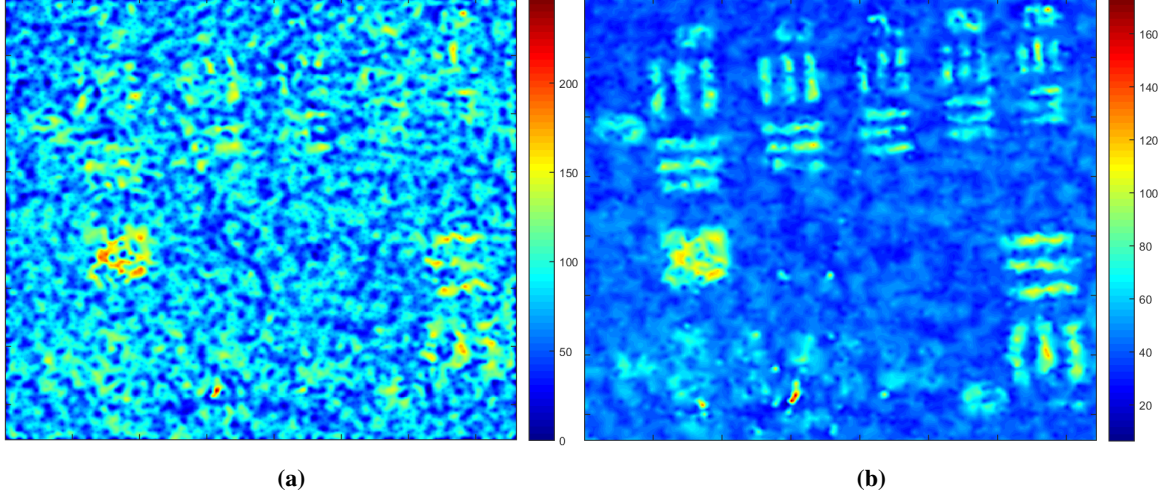


Fig. 17: (a) The Ilo algorithm output image for $\lambda = 1 \times 10^{-5}$. (b) An example of the CIo and CCo ($\lambda = 1 \times 10^{-5}$), and CCs output images. Linear interpolation is considered for CCo and CCs.

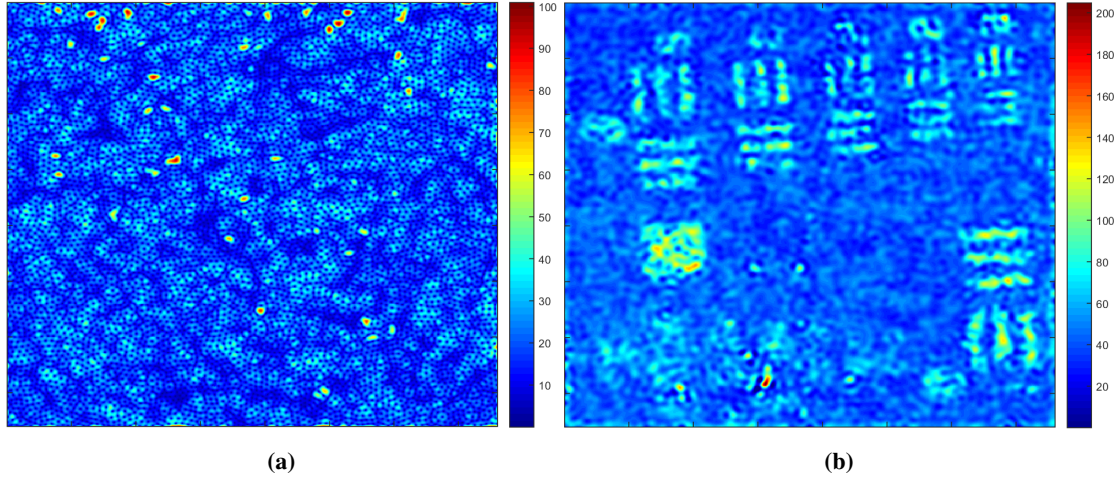


Fig. 18: (a) Non-linear interpolation by GPR of deconvolved cores of the CCo method, (b) the corresponding confidence interval image.

A. 1951 USAF resolution test chart

The 1951 USAF chart is a resolution test pattern set by US Air Force in 1951. It is widely accepted to test the resolution of optical imaging systems such as microscopes, cameras and image scanners [34]. Fig. 15(a) shows the original USAF resolution test chart used in the project. The resulting image obtained by fiber bundle is shown in Fig.15(b) which is of size 760×760 and is composed of 7776 fiber cores, a linear and non-linear (based on GPR) interpolations of central cores are presented in Fig.16. We can observe the blurring which is caused by the cross coupling effect as well as the sparsity of the data. The idea now is to

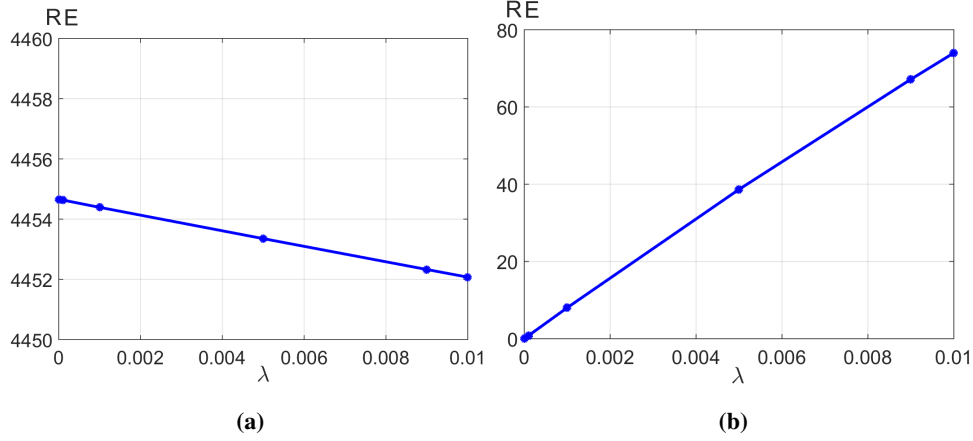


Fig. 19: Plot of RE versus regularization parameter λ for (a) Ilo, (b) CIo, and CCo.

compensate for fiber core cross coupling.

1) *Ilo*: The result of the Ilo algorithm corresponding to $\lambda = 1 \times 10^{-5}$ is presented in Fig. 17(a). The size of the Gaussian blurring kernel associated with fiber core size is fixed to 11×11 and $\sigma_C^2 = 2.2$. We can observe that the USAF pattern is not resolved at all. This is most probably because of the variation in core sizes. Fig. 19a shows a plot of RE versus different regularization parameters λ . We can observe that there is a small variation in RE for different λ in addition to its high value.

2) *CIo*, *CCo*, and *CCs*: The output of the CIo, CCo, and CCs algorithms is very similar. Fig. 17(b) shows an example of one of the output images using linear interpolation. The set of bigger bars is now better resolved and the overlap between them is reduced. The small set of bars which is at the bottom could not be resolved, which gives an indication about the resolving resolution of our endomicroscopy system. Fig. 18(b) shows non-linear interpolation of deconvolved cores using GPR. It is showing similar results to linear interpolation, however it is providing more information about the confidence interval of each pixel as shown in Fig. 18(a). We can observe that as we go away from central cores, the confidence of the interpolated intensities decreases and vice versa. A plot of RE versus different regularization parameters λ for the CIo and CCo methods is illustrated in Fig. 19(b). We can observe that there is a linear relationship between RE and λ . We also noticed that as λ increases, the deconvolved image becomes more smooth. For the CCs method, a chain of 1500 including 500 iterations as burn period is considered. The RE of this method is $RE = 5.7$.

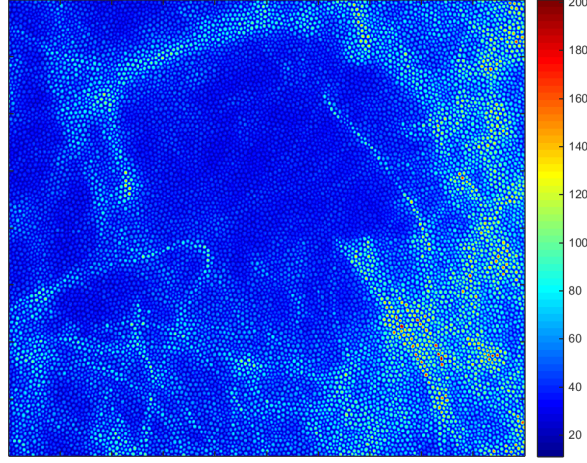


Fig. 20: Ex-vivo lung tissue imaged by an actual endomicroscopy system.

B. Ex-vivo human lung tissues

Ex-vivo measurements of human lung tissues were collected using our OEM system [4]. Fig. 20 shows the output image of the system. It is of size 1000×800 and is composed of 13343 fiber cores. a linear and non-linear (based on GPR) interpolations of central cores are presented in Fig.21. Similar to the USAF resolution test chart, we aim at reducing cross coupling effect as well as getting a more resolved image.

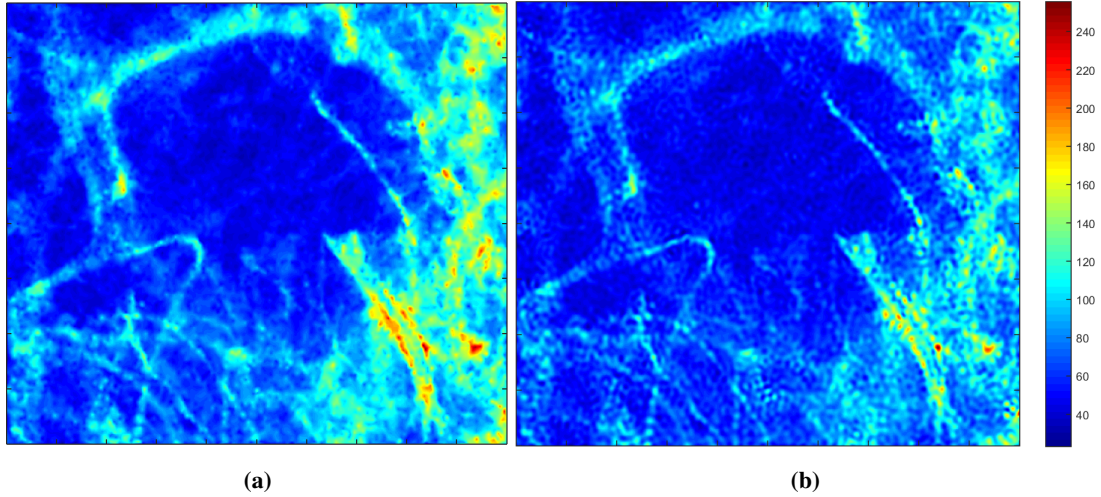


Fig. 21: (a) Linear, and (b) non-linear interpolation of central cores before deconvolution.

1) *Ilo*: Fig. 22(a) shows the output of the *Ilo* algorithm corresponding to $\lambda = 0.1$. Although the pattern is identified, it is not resolved. This is most probably because of the variation in core sizes as we have already mentioned in the USAF pattern. Fig. 24(a) shows a plot of RE

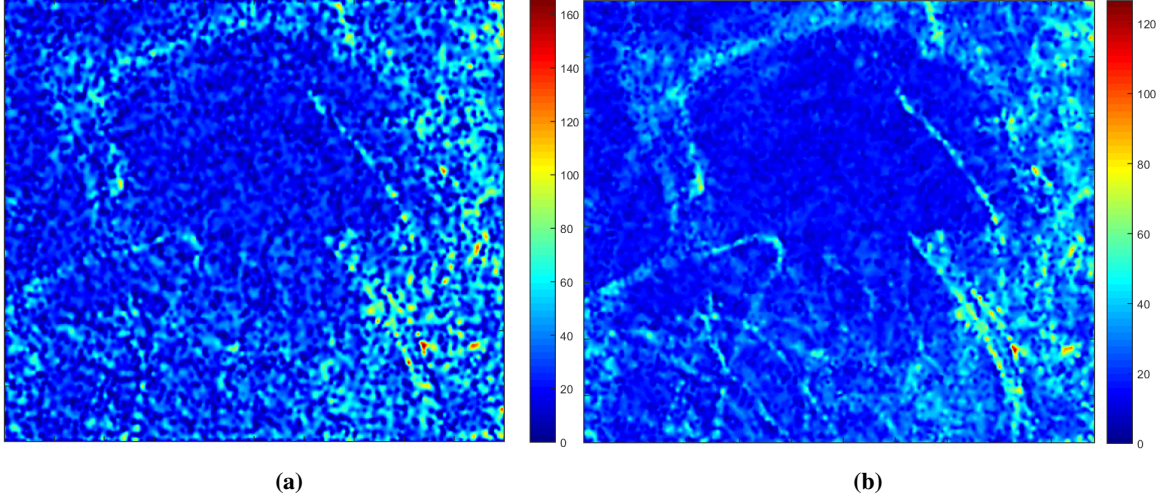


Fig. 22: (a) The Ilo algorithm output image for $\lambda = 0.1$. (b) An example of the CIo and CCo ($\lambda = 1 \times 10^{-5}$), and CCs output images. Linear interpolation is considered for CCo and CCs.

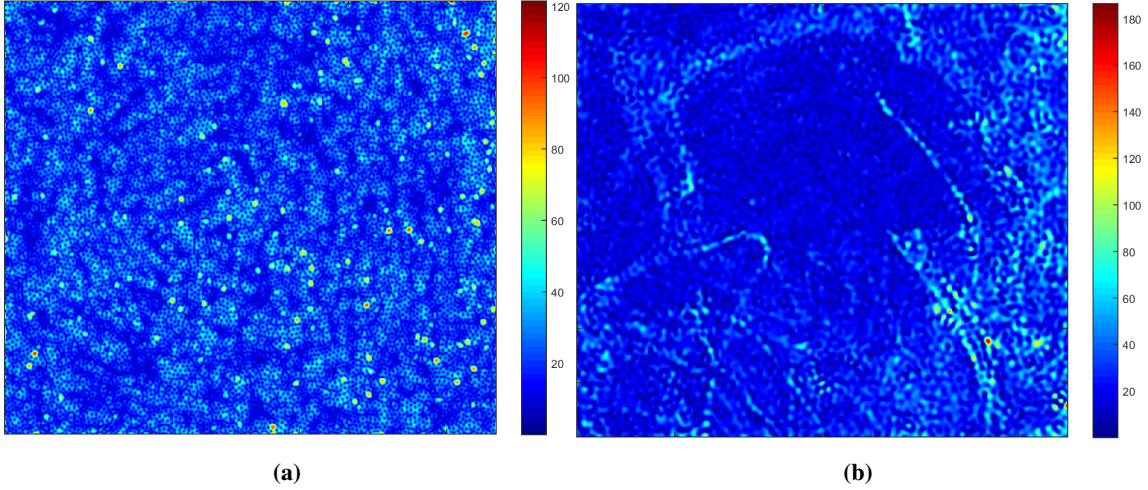


Fig. 23: (a) Non-linear interpolation by GPR of deconvolved cores of the CCo method, (b) the corresponding confidence interval image.

versus different regularization parameters λ . We can observe that there is fairly a constant relationship between RE and λ in addition to the high RE.

2) *CIo*, *CCo*, and *CCs*: Similar to the USAF resolution test chart results, the output of the CIo, CCo, and CCs algorithms is very similar. Fig. 22(b) shows an example of one of the output images using linear interpolation. Lung structure is now resolved and more identified than before deconvolution. Non-linear interpolation using GPR is shown in Fig. 23(b) which gives a more refined structures than linear interpolation. Furthermore, GPR provides more information about confidence intervals as shown in Fig. 23(a). We can observe that as we

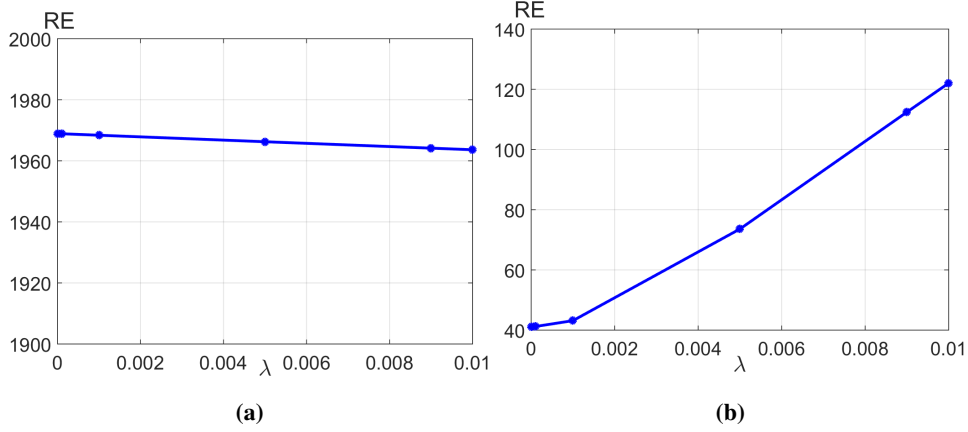


Fig. 24: Plot of RE versus regularization parameter λ for (a) IIo, (b) CCo, and CCo.

go away from central cores, the confidence of the interpolated intensities decreases and vice versa. A plot of RE versus different regularization parameters λ for the CCo and CCo methods is illustrated in Fig. 24(b). We can observe that there is a linear relationship between RE and λ . We also noticed that as λ increases, the deconvolved image becomes more smooth. For the CCs method, a chain of 1500 including 500 iterations as burn period is considered.

VI. CONCLUSION

This work investigated three optimization and one simulation based algorithms to deconvolve and restore optical endomicroscopy images. Different approaches have been considered. One might want to reconstruct the full image directly or only estimate the intensities at fiber core locations, in this case the estimated full image can be recovered by interpolating the deconvolved subsampled image. Extensive simulations conducted on both synthetic and real datasets showed good performance for the proposed methods. All of the methods showed a fairly linear relationship between RMSE and noise variance for a given cross coupling level, and a constant relationship between RMSE and cross coupling levels at constant noise variance. We observed a trade off between the deconvolution and restoration quality and the computation cost of the algorithms. The IIo method showed the lowest RMSE compared to the rest of the methods, but is the most computationally expensive method. The CCo method showed the lowest computation cost with a reasonable performance degradation. As a future work, we propose to consider temporal information while deconvolving, consider different deconvolution kernels based on core size, test different priors for the CCs method, and compare the performance of bacteria detection on raw and deconvolved data.

APPENDIX

In this appendix, we detail the ADMM-based deconvolution via two deconvolution operators (IIo) introduced in subsection III-A. In order to solve for $\mathbf{u}, \mathbf{v}_1, \dots, \mathbf{v}_7$, we construct the (AL) corresponding to (8) as follows

$$\begin{aligned} \mathcal{L}(\mathbf{u}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6, \mathbf{v}_7, \mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4, \mathbf{d}_5, \mathbf{d}_6, \mathbf{d}_7) = & \frac{1}{2} \|\mathbf{v}_1 - \mathbf{y}\|_2^2 + \lambda \phi(\mathbf{v}_6) + i_{\mathbb{R}^+}(\mathbf{v}_7) \\ & + \frac{\mu}{2} \|\mathbf{C}\mathbf{v}_2 - \mathbf{v}_1 - \mathbf{d}_1\|_2^2 + \frac{\mu}{2} \|\mathbf{D}^T \mathbf{v}_3 - \mathbf{v}_2 - \mathbf{d}_2\|_2^2 + \frac{\mu}{2} \|\mathbf{H}\mathbf{v}_4 - \mathbf{v}_3 - \mathbf{d}_3\|_2^2 + \\ & \frac{\mu}{2} \|\mathbf{D}\mathbf{v}_5 - \mathbf{v}_4 - \mathbf{d}_4\|_2^2 + \frac{\mu}{2} \|\mathbf{u} - \mathbf{v}_5 - \mathbf{d}_5\|_2^2 + \frac{\mu}{2} \|\mathbf{M}\mathbf{u} - \mathbf{v}_6 - \mathbf{d}_6\|_2^2 + \frac{\mu}{2} \|\mathbf{u} - \mathbf{v}_7 - \mathbf{d}_7\|_2^2 \end{aligned} \quad (33)$$

where $\mathbf{d}_1, \dots, \mathbf{d}_7$ are the scaled Lagrange multipliers associated to the constraints. The ADMM pseudocode corresponding to formulation (33) (and also (6), (7) and (8)) is shown in Algorithm 3.

Algorithm 3 ADMM based deconvolution via two deconvolution operators (IIo)

- 1: set $k = 0$, choose $\mu > 0$, $\mathbf{u}^{(0)}, \mathbf{v}_1^{(0)}, \dots, \mathbf{v}_7^{(0)}, \mathbf{d}_1^{(0)}, \dots, \mathbf{d}_7^{(0)}$
 - 2: **repeat**
 - 3: $\mathbf{u}^{(k+1)} \leftarrow \operatorname{argmin}_{\mathbf{u}} \mathcal{L}(\mathbf{u}, \mathbf{v}_1^{(k)}, \dots, \mathbf{v}_7^{(k)}, \mathbf{d}_1^{(k)}, \dots, \mathbf{d}_7^{(k)})$
 - 4: **for** $i = 1, \dots, 7$ **do**
 - 5: $\mathbf{v}_i^{(k+1)} \leftarrow \operatorname{argmin}_{\mathbf{v}_i} \mathcal{L}(\mathbf{u}^{(k+1)}, \mathbf{v}_1^{(k)}, \dots, \mathbf{v}_i, \dots, \mathbf{v}_7^{(k)}, \mathbf{d}_1^{(k)}, \dots, \mathbf{d}_7^{(k)})$
 - 6: **end for**
 - 7: **Update Lagrange multipliers:**
 - $\mathbf{d}_1^{(k+1)} = \mathbf{d}_1^{(k)} - \left(\mathbf{C}\mathbf{v}_2^{(k+1)} - \mathbf{v}_1^{(k+1)} \right)$
 - $\mathbf{d}_2^{(k+1)} = \mathbf{d}_2^{(k)} - \left(\mathbf{D}^T \mathbf{v}_3^{(k+1)} - \mathbf{v}_2^{(k+1)} \right)$
 - $\mathbf{d}_3^{(k+1)} = \mathbf{d}_3^{(k)} - \left(\mathbf{H}\mathbf{v}_4^{(k+1)} - \mathbf{v}_3^{(k+1)} \right)$
 - $\mathbf{d}_4^{(k+1)} = \mathbf{d}_4^{(k)} - \left(\mathbf{D}\mathbf{v}_5^{(k+1)} - \mathbf{v}_4^{(k+1)} \right)$
 - $\mathbf{d}_5^{(k+1)} = \mathbf{d}_5^{(k)} - \left(\mathbf{u}^{(k+1)} - \mathbf{v}_5^{(k+1)} \right)$
 - $\mathbf{d}_6^{(k+1)} = \mathbf{d}_6^{(k)} - \left(\mathbf{M}\mathbf{u}^{(k+1)} - \mathbf{v}_6^{(k+1)} \right)$
 - $\mathbf{d}_7^{(k+1)} = \mathbf{d}_7^{(k)} - \left(\mathbf{u}^{(k+1)} - \mathbf{v}_7^{(k+1)} \right)$
 - 8: **Update iteration** $k \leftarrow k + 1$
 - 9: **until** some stopping criterion is satisfied.
-

Algorithm 3 sequentially optimizes \mathcal{L} with respect to \mathbf{u} (step 3) and $\mathbf{v}_1, \dots, \mathbf{v}_7$ (step 5),

and then updates the Lagrange multipliers (steps 8:14). The algorithm terminates when a stopping criterion is satisfied.

We now give details about step 3 of Algorithm 3. The goal of this step is to determine the value of the variable \mathbf{u} at each iteration. Given that we run an optimization over the variable \mathbf{u} , the terms of the objective function (8) which don't contain this variable are not taken into account. The reduced optimization function becomes, then

$$\mathbf{u}^{(k+1)} \leftarrow \arg \min_{\mathbf{u}} \frac{\mu}{2} \left\| \mathbf{u} - \mathbf{v}_5^{(k)} - \mathbf{d}_5^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{M}\mathbf{u} - \mathbf{v}_6^{(k)} - \mathbf{d}_6^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{u} - \mathbf{v}_7^{(k)} - \mathbf{d}_7^{(k)} \right\|_2^2 \quad (34)$$

It is easy to show that the solution of this problem is

$$\mathbf{u}^{(k+1)} \leftarrow (\mathbf{M}^T \mathbf{M} + 2\mathbf{I})^{-1} \left(\mathbf{M}^T \mathbf{b}_6^{(k)} + \mathbf{b}_5^{(k)} + \mathbf{b}_7^{(k)} \right) \quad (35)$$

where \mathbf{I} is the identity matrix, \mathbf{M}^T represents the transpose of \mathbf{M} and: $\mathbf{b}_5^{(k)} = \mathbf{v}_5^{(k)} + \mathbf{d}_5^{(k)}$, $\mathbf{b}_6^{(k)} = \mathbf{v}_6^{(k)} + \mathbf{d}_6^{(k)}$ and $\mathbf{b}_7^{(k)} = \mathbf{v}_7^{(k)} + \mathbf{d}_7^{(k)}$.

On the other hand, step 5 of the of Algorithm 3 computes the values of the variables $\mathbf{v}_1, \dots, \mathbf{v}_7$ at the current iteration. In order to compute \mathbf{v}_1 , the optimization problem to solve is

$$\mathbf{v}_1^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_1} \frac{1}{2} \left\| \mathbf{v}_1^{(k)} - \mathbf{y} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{C}\mathbf{v}_2^{(k)} - \mathbf{v}_1 - \mathbf{d}_1^{(k)} \right\|_2^2 \quad (36)$$

whose solution can be easily shown as

$$\mathbf{v}_1^{(k+1)} \leftarrow \frac{1}{\mu + 1} \left[\mathbf{y} + \mu \left(\mathbf{C}\mathbf{v}_2^{(k)} - \mathbf{d}_1^{(k)} \right) \right] \quad (37)$$

To compute \mathbf{v}_2 , the optimization problem to be solved is

$$\mathbf{v}_2^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_2} \frac{\mu}{2} \left\| \mathbf{C}\mathbf{v}_2 - \mathbf{v}_1^{(k+1)} - \mathbf{d}_1^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{D}^T \mathbf{v}_3^{(k)} - \mathbf{v}_2 - \mathbf{d}_2^{(k)} \right\|_2^2 \quad (38)$$

whose solution is

$$\mathbf{v}_2^{(k+1)} \leftarrow (\mathbf{C}^H \mathbf{C} + \mathbf{I})^{-1} \left(\mathbf{C}^H \mathbf{b}_1^{(k+1)} + \mathbf{D}^T \mathbf{v}_3^{(k)} - \mathbf{d}_2^{(k)} \right) \quad (39)$$

where $\mathbf{b}_1^{(k+1)} = \mathbf{v}_1^{(k+1)} + \mathbf{d}_1^{(k)}$.

Using a single core size to represent \mathbf{C} allows it to be a block circulant matrix with circulant blocks, that can be factorized as $\mathbf{C} = \mathbf{R}^H \mathbf{P} \mathbf{R}$, where \mathbf{R} is the unitary matrix ($\mathbf{R}^H = \mathbf{R}^{-1}$) representing the Discrete Fourier Transform (DFT) and \mathbf{P} is diagonal, which can make the computation faster, thus

$$(\mathbf{C}^H \mathbf{C} + \mathbf{I})^{-1} = \mathbf{R}^H (|\mathbf{P}|^2 + \mathbf{I})^{-1} \mathbf{R}$$

where $|P|^2$ is the matrix with the squared absolute values of the entries of P .

In order to compute \mathbf{v}_3 , the optimization problem to solve is

$$\mathbf{v}_3^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_3} \frac{\mu}{2} \left\| \mathbf{D}^T \mathbf{v}_3 - \mathbf{v}_2^{(k+1)} - \mathbf{d}_2^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{H} \mathbf{v}_4^{(k)} - \mathbf{v}_3 - \mathbf{d}_3^{(k)} \right\|_2^2 \quad (40)$$

whose solution is

$$\mathbf{v}_3^{(k+1)} \leftarrow (\mathbf{D} \mathbf{D}^T + \mathbf{I})^{-1} \left(\mathbf{D} \mathbf{b}_2^{(k+1)} + \mathbf{H} \mathbf{v}_4^{(k)} - \mathbf{d}_3^{(k)} \right) \quad (41)$$

where $\mathbf{b}_2^{(k+1)} = \mathbf{v}_2^{(k+1)} + \mathbf{d}_2^{(k)}$.

To compute \mathbf{v}_4 , the optimization problem to be solved is

$$\mathbf{v}_4^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_4} \frac{\mu}{2} \left\| \mathbf{H} \mathbf{v}_4 - \mathbf{v}_3^{(k+1)} - \mathbf{d}_3^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{D} \mathbf{v}_5^{(k)} - \mathbf{v}_4 - \mathbf{d}_4^{(k)} \right\|_2^2 \quad (42)$$

whose solution is

$$\mathbf{v}_4^{(k+1)} \leftarrow (\mathbf{H}^T \mathbf{H} + \mathbf{I})^{-1} \left(\mathbf{H}^T \mathbf{b}_3^{(k+1)} + \mathbf{D} \mathbf{v}_5^{(k)} - \mathbf{d}_4^{(k)} \right) \quad (43)$$

where $\mathbf{b}_3^{(k+1)} = \mathbf{v}_3^{(k+1)} + \mathbf{d}_3^{(k)}$.

In order to compute \mathbf{v}_5 , we solve the optimization problem

$$\mathbf{v}_5^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_5} \frac{\mu}{2} \left\| \mathbf{D} \mathbf{v}_5 - \mathbf{v}_4^{(k+1)} - \mathbf{d}_4^{(k)} \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{u}^{(k+1)} - \mathbf{v}_5 - \mathbf{d}_5^{(k)} \right\|_2^2 \quad (44)$$

having the solution

$$\mathbf{v}_5^{(k+1)} \leftarrow (\mathbf{D}^T \mathbf{D} + \mathbf{I})^{-1} \left(\mathbf{D}^T \mathbf{b}_4^{(k+1)} + \mathbf{u}^{(k+1)} - \mathbf{d}_5^{(k)} \right) \quad (45)$$

where $\mathbf{b}_4^{(k+1)} = \mathbf{v}_4^{(k+1)} + \mathbf{d}_4^{(k)}$.

To compute \mathbf{v}_6 , we solve the optimization problem

$$\mathbf{v}_6^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_6} \lambda \phi(\mathbf{v}_6) + \frac{\mu}{2} \left\| \mathbf{M} \mathbf{u}^{(k+1)} - \mathbf{v}_6 - \mathbf{d}_6^{(k)} \right\|_2^2 \quad (46)$$

whose solution is

$$\mathbf{v}_6^{(k+1)} \leftarrow \frac{\mu}{\mu + 2\lambda} \left(\mathbf{M} \mathbf{u}^{(k+1)} - \mathbf{d}_6^{(k)} \right) \quad (47)$$

In order to compute \mathbf{v}_7 , we solve the optimization problem

$$\mathbf{v}_7^{(k+1)} \leftarrow \arg \min_{\mathbf{v}_7} i_{\mathbb{R}^+}(\mathbf{v}_7) + \frac{\mu}{2} \left\| \mathbf{u}^{(k+1)} - \mathbf{v}_7 - \mathbf{d}_7^{(k)} \right\|_2^2 \quad (48)$$

The role of $i_{\mathbb{R}^+}$ term is to project the solution onto the non-negative orthant, the value of \mathbf{v}_7 is given by

$$\mathbf{v}_7^{(k+1)} \leftarrow \max \left(\mathbf{u}^{(k+1)} - \mathbf{d}_7^{(k)}, \mathbf{0} \right) \quad (49)$$

ACKNOWLEDGMENT

We would like to thank Engineering and Physical Sciences Research Council (EPSRC, United Kingdom) Interdisciplinary Research Collaboration grant EP/K03197X/1 for funding this work.

REFERENCES

- [1] J. Chastre and J.-Y. Fagon, “Ventilator-associated pneumonia,” American journal of respiratory and critical care medicine, vol. 165, no. 7, pp. 867–903, 2002.
- [2] P. Johnston, D. F. McAuley, and C. M. O’Kane, “Novel pulmonary biomarkers in the diagnosis of vap,” Thorax, vol. 65, no. 3, pp. 190–192, 2010.
- [3] V. S. Baselski and R. G. Wunderink, “Bronchoscopic diagnosis of pneumonia,” Clinical microbiology reviews, vol. 7, no. 4, pp. 533–558, 1994.
- [4] N. Krstajić, A. R. Akram, T. R. Choudhary, N. McDonald, M. G. Tanner, E. Pedretti, P. A. Dalgarno, E. Scholefield, J. M. Girkin, A. Moore et al., “Two-color widefield fluorescence microendoscopy enables multiplexed molecular imaging in the alveolar space of human lung tissue,” Journal of Biomedical Optics, vol. 21, no. 4, pp. 046 009–046 009, 2016.
- [5] A. R. Akram, N. Avlonitis, A. Lilienkampf, A. M. Perez-Lopez, N. McDonald, S. V. Chankeshwara, E. Scholefield, C. Haslett, M. Bradley, and K. Dhaliwal, “A labelled-ubiquicidin antimicrobial peptide for immediate in situ optical detection of live bacteria in human alveolar lung tissue,” Chemical Science, vol. 6, no. 12, pp. 6971–6979, 2015.
- [6] N. Avlonitis, M. Deburne, T. Aslam, N. McDonald, C. Haslett, K. Dhaliwal, and M. Bradley, “Highly specific, multi-branched fluorescent reporters for analysis of human neutrophil elastase,” Organic & biomolecular chemistry, vol. 11, no. 26, pp. 4414–4418, 2013.
- [7] T. Aslam, A. Miele, S. V. Chankeshwara, A. Megia-Fernandez, C. Michels, A. R. Akram, N. McDonald, N. Hirani, C. Haslett, M. Bradley et al., “Optical molecular imaging of lysyl oxidase activity–detection of active fibrogenesis in human lung tissue,” Chemical Science, vol. 6, no. 8, pp. 4946–4953, 2015.
- [8] G. Le Goualher, A. Perchant, M. Genet, C. Cavé, B. Viellerobe, F. Berier, B. Abrat, and N. Ayache, “Towards optical biopsies with an integrated fibered confocal fluorescence microscope,” in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004. Springer, 2004, pp. 761–768.
- [9] N. Ayache, T. Vercauteren, G. Malandain, F. Oberrietter, N. Savoie, and A. Perchant, “Processing and mosaicing of fibered confocal images,” in MICCAI Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB06), 2006.
- [10] P. McCool, Y. Altmann, A. Perperidis, and S. McLaughlin, “Robust markov random field outlier detection and removal in subsampled images,” in IEEE workshop on Statistical Signal Processing. IEEE, 2016.
- [11] M. Bertero and P. Boccacci, Introduction to inverse problems in imaging. CRC press, 1998.
- [12] F. Šroubek and J. Flusser, “Multichannel blind iterative image restoration,” Image Processing, IEEE Transactions on, vol. 12, no. 9, pp. 1094–1106, 2003.
- [13] M. A. Figueiredo, J. M. Bioucas-Dias, and M. V. Afonso, “Fast frame-based image deconvolution using variable splitting and constrained optimization,” in Statistical Signal Processing, 2009. SSP’09. IEEE/SP 15th Workshop on. IEEE, 2009, pp. 109–112.

- [14] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, “Non-cyclic deconvolution using an augmented lagrangian method,” in EUROCON-International Conference on Computer as a Tool (EUROCON), 2011 IEEE. IEEE, 2011, pp. 1–4.
- [15] ———, “Fast image recovery using variable splitting and constrained optimization,” Image Processing, IEEE Transactions on, vol. 19, no. 9, pp. 2345–2356, 2010.
- [16] R. Courant et al., “Variational methods for the solution of problems of equilibrium and vibrations,” Bull. Amer. Math. Soc., vol. 49, no. 1, pp. 1–23, 1943.
- [17] Y. Wang, J. Yang, W. Yin, and Y. Zhang, “A new alternating minimization algorithm for total variation image reconstruction,” SIAM Journal on Imaging Sciences, vol. 1, no. 3, pp. 248–272, 2008.
- [18] J. Nocedal and S. Wright, Numerical optimization. Springer Science & Business Media, 2006.
- [19] D. Gabay and B. Mercier, “A dual algorithm for the solution of nonlinear variational problems via finite element approximation,” Computers & Mathematics with Applications, vol. 2, no. 1, pp. 17–40, 1976.
- [20] J. Eckstein and D. P. Bertsekas, “On the douglasrachford splitting method and the proximal point algorithm for maximal monotone operators,” Mathematical Programming, vol. 55, no. 1-3, pp. 293–318, 1992.
- [21] A. Wong, X. Y. Wang, and M. Gorbet, “Bayesian-based deconvolution fluorescence microscopy using dynamically updated nonstationary expectation estimates,” Scientific reports, vol. 5, 2015.
- [22] T. E. Bishop, R. Molina, and J. R. Hopgood, “Blind restoration of blurred photographs via ar modelling and mcmc,” in Image Processing, 2008. ICIIP 2008. 15th IEEE International Conference on. IEEE, 2008, pp. 669–672.
- [23] P. Ruiz, X. Zhou, J. Mateos, R. Molina, and A. K. Katsaggelos, “Variational bayesian blind image deconvolution: A review,” Digital Signal Processing, vol. 47, pp. 116–127, 2015.
- [24] F. Al-Awadhi, C. Jennison, and M. Hurn, “Statistical image analysis for a confocal microscopy two-dimensional section of cartilage growth,” Journal of the Royal Statistical Society: Series C (Applied Statistics), vol. 53, no. 1, pp. 31–49, 2004.
- [25] F. Al-Awadhi, M. Hurn, and C. Jennison, “Three-dimensional bayesian image analysis and confocal microscopy,” Journal of Applied Statistics, vol. 38, no. 1, pp. 29–46, 2011.
- [26] T. F. Chan and J. J. Shen, Image processing and analysis: variational, PDE, wavelet, and stochastic methods. Siam, 2005.
- [27] R. M. Haralock and L. G. Shapiro, Computer and robot vision. Addison-Wesley Longman Publishing Co., Inc., 1991.
- [28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends® in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.
- [29] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems,” Image Processing, IEEE Transactions on, vol. 20, no. 3, pp. 681–695, 2011.
- [30] C. Robert and G. Casella, Monte Carlo statistical methods. Springer Science & Business Media, 2013.
- [31] A. Pakman and L. Paninski, “Exact hamiltonian monte carlo for truncated multivariate gaussians,” Journal of Computational and Graphical Statistics, vol. 23, no. 2, pp. 518–542, 2014.
- [32] M. X. Goemans and D. P. Williamson, “The primal-dual method for approximation algorithms and its application to network design problems,” Approximation algorithms for NP-hard problems, pp. 144–191, 1997.
- [33] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, “Nonlinear spectral unmixing of hyperspectral images using gaussian processes,” IEEE Transactions on Signal Processing, vol. 61, no. 10, pp. 2442–2453, 2013.

- [34] SilverFast, “Silverfast resolution target (usaf 1951) by lasersoft imaging,” available online <http://www.silverfast.com/show/resolution-target/en.html>, 2012.